

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OK nucleic - nucleic search, using sw model

Run on: June 21, 2003, 19:16:44 ; Search time 1433 Seconds
(without alignments)
214.734 Million cell updates/sec

Title: US-09-581-500B-14

Perfect score: 19

Sequence: 1 cgcctcgtatcctcctcctcgt 19

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

EST:*
1: em_estbta:*
2: em_esthum:*
3: em_estina:*
4: em_estmu:*
5: em_estrov:*
6: em_estcp1:*
7: em_estcro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_man:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	17.4	91.6	522	12	BG347460
2	17.4	91.6	1039	9	AU050798
3	17.4	91.6	1101	17	CNS04VNO
4	17	89.5	187	17	AZ248177
5	17	89.5	250	17	AZ083738
6	17	89.5	772	12	BF627795

C	7	16.4	86.3	256	13	BT293206	BT293206
	8	16.4	86.3	259	9	AA884860	AA884860
	9	16.4	86.3	267	12	BF551260	BF551260
	10	16.4	86.3	292	14	BQ237385	BQ237385
	11	16.4	86.3	318	17	AZ774858	AZ774858
	12	16.4	86.3	331	10	BM125791	BM125791
	13	16.4	86.3	353	13	BG994801	BG994801
	14	16.4	86.3	364	14	TI1530	TI1530
	15	16.4	86.3	392	12	BF086663	BF086663
	16	16.4	86.3	435	9	AA998619	AA998619
	17	16.4	86.3	465	9	AA998179	AA998179
	18	16.4	86.3	466	9	AA996405	AA996405
	19	16.4	86.3	497	12	BF410904	BF410904
	20	16.4	86.3	499	17	AZ127320	AZ127320
	21	16.4	86.3	516	10	BE374792	BE374792
	22	16.4	86.3	525	12	BF190964	BF190964
	23	16.4	86.3	528	9	AI178165	AI178165
	24	16.4	86.3	535	10	BT201299	BT201299
	25	16.4	86.3	538	13	BZ744737	BZ744737
	26	16.4	86.3	538	17	AZ644006	AZ644006
	27	16.4	86.3	545	17	AQ041975	AQ041975
	28	16.4	86.3	554	10	BE100774	BE100774
	29	16.4	86.3	577	17	AQ870513	AQ870513
	30	16.4	86.3	578	17	AZ483135	AZ483135
	31	16.4	86.3	591	17	AQ491443	AQ491443
	32	16.4	86.3	616	12	BG162703	BG162703
	33	16.4	86.3	685	10	BE410269	BE410269
	34	16.4	86.3	705	14	BQ209524	BQ209524
	35	16.4	86.3	719	17	AZ848914	AZ848914
	36	16.4	86.3	720	12	BE729865	BE729865
	37	16.4	86.3	720	13	BI650819	BI650819
	38	16.4	86.3	748	12	BG671778	BG671778
	39	16.4	86.3	782	13	BI558945	BI558945
	40	16.4	86.3	831	12	BE876298	BE876298
	41	16.4	86.3	857	9	AU141539	AU141539
	42	16.4	86.3	884	17	AQ742984	AQ742984
	43	16.4	86.3	1003	17	AG110540	AG110540
	44	16.4	86.3	2307	11	BC021805	BC021805
	45	16.4	86.3	2586	11	BC007638	BC007638

ALIGNMENTS

RESULT 1	BG347460	522 bp	mRNA	linear	EST 28-FRB-2001
LOCUS	BG347460				
DEFINITION	dab87a11.x1 NICHD XCC Emb2 Xenopus laevis cDNA clone IMAGE:4404381				
	3 similar to SW:NOP4_YEAST P37838 NUCLEOLAR PROTEIN NOP4 ; , mRNA				
ACCESSION	BG347460				
VERSION	BG347460.1	GI:13167884			
KEYWORDS	EST.				
SOURCE	African clawed frog.				
ORGANISM	Xenopus laevis				
REFERENCE	1 (bases 1 to 522)				
AUTHORS	Clifton,S., Johnson,S.L., Blumberg,B., Song,J., Hillier,L., Page,D., Martin,J., Wylie,T., Underwood,K., Theising,B., Bowers,Y., Person,B., Gibbons M., Harvey,N., Ritter,E., Jackson,Y., McCann,R., Materson,R. and Wilson,R.				
	Washu Xenopus EST project, 1999				
TITLE	Unpublished (1999)				
COMMENT	Contact: Sandy Clifton, Ph.D. Washu Xenopus EST project, 1999 Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: estewatson.wustl.edu Library constructed by Life Technologies. DNA Sequencing by:				

JOURNAL
COMMENT

Unpublished (1999)
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong

(pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>) or from Resea ch Genetics (<http://info@resgen.com>). BAC end page: http://www.tigr.org/tldb/bac_ends/mouse/bac_end_intro.html

Plate: 39 row: M column: 14

Seq primer: T7
Class: BAC ends.

FEATURES
source

Location/Qualifiers
1..187
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-39M14"
/clone_id="RPCI-23"
/sex="Female"
/lab_host="DH10B"
/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site 1: EcoRI; Site 2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACe3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."

BASE COUNT
ORIGIN

58 a 30 c 65 g 34 t

Query Match 89.5%; Score 17; DB 17; Length 187;
Best Local Similarity 100.0%; Pred. No. 2e+03;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GCTCTGATTCCTGCTCT 18
|||||
Db 27 GCTCTGATTCCTGCTCT 43

RESULT 5
AZ083738 250 bp DNA linear GSS 08-MAY-2000
LOCUS
DEFINITION
RPCI-23-22N10.TV RPCI-23 Mus musculus genomic clone RPCI-23-22N10,
DNA sequence.
ACCESSION
AZ083738
VERSION
AZ083738.1 GI:7725471
KEYWORDS
GSS.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 250)
Bao, S., Nieman, W., Feldblum, T., Malek, J., Shatsman, S., Akhurst,
B., Levins, M., Moggan, S., Tesgaye, G., Geer, K., Krol, M., de Jong, P.
and Fraser, C.M.
Mouse BAC End Sequences from Library RPCI-23
Unpublished (1999)
Other GSSs: RPCI-23-22N10.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong

TITLE
JOURNAL
COMMENT

Mouse BAC End Sequences from Library RPCI-23
Unpublished (1999)
Other GSSs: RPCI-23-22N10.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong

FEATURES
source

(pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>) or from Resea ch Genetics (<http://info@resgen.com>). BAC end page: http://www.tigr.org/tldb/bac_ends/mouse/bac_end_intro.html

Plate: 22 row: N column: 10

Seq primer: SP6

Class: BAC ends.

FEATURES
source

Location/Qualifiers
1..250
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-22N10"
/clone_id="RPCI-23"
/sex="Female"
/lab_host="DH10B"
/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site 1: EcoRI; Site 2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACe3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."

BASE COUNT
ORIGIN

71 a 35 c 80 g 64 t

Query Match 89.5%; Score 17; DB 17; Length 250;
Best Local Similarity 100.0%; Pred. No. 2.2e+03;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GCTCTGATTCCTGCTCT 18
|||||
Db 44 GCTCTGATTCCTGCTCT 60

RESULT 6
BF627795/c 772 bp mRNA linear EST 22-OCT-2001
LOCUS
DEFINITION
HVSMB0005N04f Hordeum vulgare seedling shoot EST library
HYCDNA0002 (Dehydration stress) Hordeum vulgare cDNA clone
HVSMB0005N04f, mRNA sequence.
ACCESSION
BF627795
VERSION
BF627795.2 GI:13090001
KEYWORDS
EST.
SOURCE
Hordeum vulgare.
ORGANISM
Hordeum vulgare.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
Triticeae; Hordeum.
1 (bases 1 to 772)
Wing, R., Close, T.J., Kleinhofs, A., Wise, R., Begum, D., Fritsch, D., Yu,
Y., Henry, D., Palmer, M., Rambo, T., Simons, J., Choi, D.W., Fenton,
R.D., Oates, R. and Main, D.
Development of a genetically and physically anchored EST resource
for barley genomics: Morex drought-stressed seedling shoot cDNA
library
Unpublished (2001)
On Dec 19, 2000 this sequence version replaced gi:11891953.
Contact: Wing RA
Clemson University
Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Total hg bases = 131
Seq primer: AATTACCTCACTAAGG
High quality sequence stop: 215.
location/Qualifiers
1..772
/organism="Hordeum vulgare"
/cultivar="Morex"
/db_xref="taxon:4513"

TITLE
JOURNAL
COMMENT

Unpublished (2001)
On Dec 19, 2000 this sequence version replaced gi:11891953.
Contact: Wing RA
Clemson University
Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Total hg bases = 131
Seq primer: AATTACCTCACTAAGG
High quality sequence stop: 215.
location/Qualifiers
1..772
/organism="Hordeum vulgare"
/cultivar="Morex"
/db_xref="taxon:4513"

FEATURES
source

Unpublished (2001)
On Dec 19, 2000 this sequence version replaced gi:11891953.
Contact: Wing RA
Clemson University
Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Total hg bases = 131
Seq primer: AATTACCTCACTAAGG
High quality sequence stop: 215.
location/Qualifiers
1..772
/organism="Hordeum vulgare"
/cultivar="Morex"
/db_xref="taxon:4513"

SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 259)
AUTHORS NC1-GCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
This clone is available royalty-free through LNL; contact the IMAGE Consortium (infoimage.lnl.gov) for further information. Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Seq primer: -40m3 fwd. RT from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1..259
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1467435"
/clone_lib="Soares_NFL_T_GBC_S1"
/lab_host="MDH10B"
/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung NBH19W, testis NHT, and B-cell NC1 CGAP GCBI) were mixed, and ss circles were made in vitro. Following RAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 297480-302087, 682632-687233, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 50 a 93 c 57 g 59 t
ORIGIN

Query Match 86.3%; Score 16.4; DB 9; Length 259;
Best Local Similarity 94.4%; Pred. No. 4e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GCTCTGATTCCTGCTCTG 19
DB 42 GCTCTGATTCCTGCTCTG 59

RESULT 9 267 bp mRNA linear EST 12-DEC-2000
LOCUS BF551260
DEFINITION UT-R-C0-19-e-05-0-UT.1 UT-R-C0 Rattus norvegicus cDNA clone
ACCESSION BF551260
KEYWORDS UT-R-C0-19-e-05-0-UT 5', mRNA sequence.
SOURCE RST.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 267)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565

EMAIL: msoares@blue.weeg.iowa.edu
CDNA Library Preparation: M.B. Soares Lab Clone distribution:
Clones will be available through Research Genetics (www.resgen.com)
This clone is also available through the I.M.A.G.E. Consortium at LNL (infoimage.lnl.gov). IMAGE ID= 1774202
Seq primer: M13 Forward.
Location/Qualifiers
1..267
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UT-R-C0-19-e-05-0-UT"
/clone_lib="UT-R-C0"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; The UT-R-C0 library is a subtracted library derived from the UT-R-A1 and UT-R-E1 libraries. The UT-R-A1 library consisted of a mixture of individually tagged normalized libraries constructed from rat placenta, adult lung, brain, liver, kidney, heart, spleen, ovary, and muscle. The UT-R-E1 library consisted of a mixture of individually tagged normalized libraries constructed from 8, 12 and 18-day embryo. The tag is a string of 3-5 nucleotides present between the Not I site and the oligo-dT track which allows identification of the library of origin of a clone within the mixture. The subtracted library (UT-R-C0) was constructed as follows: PCR amplified cDNA inserts from a pool of UT-R-A1 and UT-R-E1 clones from which 3' ESTs had been derived was used as a driver in a hybridization with the pooled UT-R-A1 and UT-R-E1 library in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the UT-R-C0 library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6: 791-806, 1996)."

BASE COUNT 56 a 82 c 52 g 77 t
ORIGIN

Query Match 86.3%; Score 16.4; DB 12; Length 267;
Best Local Similarity 94.4%; Pred. No. 4e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GCTCTGATTCCTGCTCTG 19
DB 219 GCTCTGATTCCTGCTCTG 236

RESULT 10 292 bp mRNA linear EST 03-MAY-2002
LOCUS B0237385
DEFINITION TA05017F05F TA05 Triticum aestivum cDNA clone TA05017F05F, mRNA sequence.
ACCESSION B0237385
KEYWORDS B0237385.1 GI:20433261
SOURCE RST.
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticaceae; Triticum.
REFERENCE 1 (bases 1 to 292)
AUTHORS Cloutier,S.
TITLE Wheat functional genomics - Glenlea developing seeds cDNA libraries
JOURNAL Unpublished (2002)
COMMENT Contact: Dr. Sylvie Cloutier
Cereal Research Centre, Agriculture and Agri-food Canada
195 Dafoe Rd, Winnipeg, MB, Canada R3T 2M5
Tel: (204) 983-2340

Fax: (204) 983-4604
Email: scloutiere@agr.ca
was cloned directionally, not all sequences generated with reverse primer were from the 5' end (same with forward primer and 3' end).
Average insert size is >2.0 kb
Plate: 017 row: F column: 05
Seq primer: M13 Forward.

FEATURES

Source
1. 292
Location/Qualifiers
/organism="Triticum aestivum"
/cultivar="Glenlea"
/db_xref="taxon:4565"
/clone="TA05017P05P"
/clone_lib="TA05"
/tissue_type="developing seeds"
/dev_stage="5 days after anthesis"
/lab_host="E. coli DH10B"
/note="Vector: pSPORT-P (Invitrogen Technologies); Site 1:
NotI; Site 2: MluI; mRNA obtained from wheat seeds of
cultivar Glenlea 5 days post-anthesis"
BASE COUNT 50 a 67 c 109 g 66 t
ORIGIN

Query Match 86.3%; Score 16.4; DB 14; Length 292;
Best Local Similarity 94.4%; Pred. No. 4.1e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
CY 2 GCTCTGATTCCTGCTCTG 19
DB 58 GCTCTGATTCCTGCTCTG 75

RESULT 11
A2774858 318 bp DNA linear GSS 16-FEB-2001
LOCUS 2M0004J1.R Mouse 10kb Plasmid UGCGM library Mus musculus genomic
DEFINITION clone UGCG2M0004J1.R, DNA sequence.
A2774858
ACCESSION A2774858.1 GI:12900749
VERSION GSS.
KEYWORDS house mouse.
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE
AUTHORS Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 318)
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausen, A.
and Wright, D., Weiss, R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: dunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0004 row: J column: 11
Seq primer: CACACAGGAACAGCTATGACC
Class plasmid ends
High quality sequence stop: 318.
Location/Qualifiers
1. 318
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UGCG2M0004J1"
/clone_lib="Mouse 10kb plasmid UGCGM library"
/sex="Male"

FEATURES

Source
1. 318
Location/Qualifiers
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UGCG2M0004J1"
/clone_lib="Mouse 10kb plasmid UGCGM library"
/sex="Male"

/lab host="E. coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: pMD42uv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (gil472114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

Query Match 86.3%; Score 16.4; DB 17; Length 318;
Best Local Similarity 94.4%; Pred. No. 4.2e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
CY 2 GCTCTGATTCCTGCTCTG 19
DB 70 GCTCTGATTCCTGCTCTG 87

RESULT 12
B125791/c 331 bp mRNA linear EST 28-JUN-2000
LOCUS B125791 RIKEN full-length enriched, 16 days neonate cerebellum Mus
DEFINITION musculus cDNA clone 963008W03 3', mRNA sequence.
B125791
ACCESSION B125791.1 GI:8780123
VERSION EST.
KEYWORDS house mouse.
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE
AUTHORS Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 331)
Komo, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T., Carninci,
P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N.,
Hirozane, T., Hori, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M.,
Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N.,
Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C., Kusakabe, M.,
Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y.,
Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata,
Y., Shigemoto, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Sugatara, Y.,
Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomioka, N., Toya,
T., Tsunoda, Y., Watanabe, A., Watanabe, S., Yamamura, T., Yamataka, I.,
Yano, R., Yasumitsu, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino,
M., Muramatsu, M. and Hayashizaki, Y.
RIKEN Mouse ESTs (Komo, H., et al.)
Unpublished (2000)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gscc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoaka, S., Sasaki,
N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Thermolabile and thermolabile enzymes by
thermalase and its application for the synthesis of full length
cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)

FEATURES
Source
1. 331
Location/Qualifiers
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UGCG2M0004J1"
/clone_lib="Mouse 10kb plasmid UGCGM library"
/sex="Male"

Source
1. 331
Location/Qualifiers
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UGCG2M0004J1"
/clone_lib="Mouse 10kb plasmid UGCGM library"
/sex="Male"

BASE COUNT	96 a	96 c	86 g	86 t	-	-
------------	------	------	------	------	---	---

ORIGIN

Query Match 86.3%; Score 16.4; DB 14; Length 364;
 Best Local Similarity 94.4%; Pred. No. 4.4e+03;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CGCTCGATTCCTGCTCT 18
 |||||
 Db 95 CGCTCGATTCCTGCTCT 112

RESULT 15

BR086663/c 392 bp mRNA linear EST 19-OCT-2000
 LOCUS CM0-GN0079-150900-547-d04 GN0079 Homo sapiens cDNA, mRNA sequence.
 DEFINITION BR086663
 ACCESSION BR086663
 VERSION BR086663.1 GI:10892373
 KEYWORDS EST.

SOURCE

ORGANISM

Homo sapiens
 human.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCES

1 (bases 1 to 392)
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
 Bunnstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
 M.J., Soares, F., Brentani, R.R., Reis, L.P., de Souza, S.J. and
 Simpson, A.U.

TITLE

JOURNAL

MEDLINE

COMMENT

Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 20202663
 Contact: Simpson A.U.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil

Tel: +55-11-2704922
 Fax: +55-11-2707001

Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?rl=kt2-CM0-GN0079-150
 900-547-d04&rl=2000-09-15&rl=41)

Seq primer: puc 18 forward
 High quality sequence start: 39
 High quality sequence stop: 392.
 Location/Qualifiers

FEATURES

source

1..392
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="GN0079"
 /dev_stage="Adult"
 /note="Organ: Placenta, normal; Vector: puc18; Site 1: SmaI
 ; Site 2: SmaI; A mini-library was made by cloning
 products derived from ORS276 PCR (U.S. Letters Patent
 application No. 196,716 - Ludwig Institute for Cancer
 Research) profiles into the pUC 18 vector. Reverse
 transcription of tissue mRNA and cDNA amplification were
 performed under low stringency conditions."

BASE COUNT

91 a 97 c 74 g 130 t

ORIGIN

Query Match 86.3%; Score 16.4; DB 12; Length 392;
 Best Local Similarity 94.4%; Pred. No. 4.5e+03;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 GGTCTGATTCCTGCTCTG 19
 |||||
 Db 39 GGTCTGATTCCTGCTCTG 22

Sequence 59, Application US/10028072
Publication No. US2003004311A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin J.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang
TITLE OF INVENTION:
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/10/028,072
CURRENT FILING DATE: 2001-12-19
PRIOR APPLICATION NUMBER: 60/049911
PRIOR FILING DATE: 1997-06-18
PRIOR APPLICATION NUMBER: 60/056974
PRIOR FILING DATE: 1997-08-26
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059115
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059117
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059122
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059184
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059352
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059588
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059836
PRIOR FILING DATE: 1997-09-24
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062285
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062287
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062814
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/062816
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063045
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063082
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/063127
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063327
PRIOR FILING DATE: 1997-10-27
PRIOR APPLICATION NUMBER: 60/063329
PRIOR FILING DATE: 1997-10-27
PRIOR APPLICATION NUMBER: 60/063550
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063561
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063704
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063733

PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063735
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063738
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063755
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064248
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/064809
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065846
PRIOR FILING DATE: 1997-11-17
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/066453
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066511
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/069212
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069278
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069334
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069694
PRIOR FILING DATE: 1997-12-16
PRIOR APPLICATION NUMBER: 60/072320
PRIOR FILING DATE: 1998-01-23
PRIOR APPLICATION NUMBER: 60/073612
PRIOR FILING DATE: 1998-02-04
PRIOR APPLICATION NUMBER: 60/074086
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/074092
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-02-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081695
PRIOR FILING DATE: 1998-04-14
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081818
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082999
PRIOR FILING DATE: 1998-04-24
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07

```

; PRIOR APPLICATION NUMBER: 60/085149
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085379
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/086414
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/086430
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088730
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088741
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090538
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07

Query Match      86.3%; Score 16.4; DB 9; Length 2550;
Best Local Similarity 94.4%; Pred. No. 80;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 GCTCTGATTCGTCTGTG 19
Db      438 GCTCTGATTCGTCTGTG 455

RESULT 3
US-10-121-049-59
; Sequence 59, Application US/10121049
; Publication No. US2003002239A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Naureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
```

```

; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Guiney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C17
; CURRENT FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: US/10/121,049
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 59
; LENGTH: 2550
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-123-904-59

Query Match      86.3%; Score 16.4; DB 9; Length 2550;
Best Local Similarity 94.4%; Pred. No. 80;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 GCTCTGATTCGTCTGTG 19
Db      438 GCTCTGATTCGTCTGTG 455

RESULT 4
US-10-123-904-59
; Sequence 59, Application US/10123904
; Publication No. US20030022328A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Naureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Guiney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C54
; CURRENT FILING DATE: 2002-04-16
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 59
; LENGTH: 2550
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-123-904-59

Query Match      86.3%; Score 16.4; DB 9; Length 2550;
Best Local Similarity 94.4%; Pred. No. 80;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

? APPLICANT: Baker, Kevin P.
 ? APPLICANT: Beresini, Maureen
 ? APPLICANT: DeForge, Laura
 ? APPLICANT: Desnoyers, Luc
 ? APPLICANT: Filvaroff, Ellen
 ? APPLICANT: Gao, Wei-Qiang
 ? APPLICANT: Gerritsen, Mary E.
 ? APPLICANT: Godard, Audrey
 ? APPLICANT: Godowski, Paul J.
 ? APPLICANT: Gurney, Austin L.
 ? APPLICANT: Sherwood, Steven
 ? APPLICANT: Smith, Victoria
 ? APPLICANT: Stewart, Timothy A.
 ? APPLICANT: Tumas, Daniel
 ? APPLICANT: Watanabe, Colin K
 ? APPLICANT: Wood, William
 ? APPLICANT: Zhang, Zemin
 ? TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEOTI

RESULT 8
US-10-176-921-59
; Sequence 59; Application US/10176922
; Publication No. US200300227276A1
; GENERAL INFORMATION:
; APPLICANT: Baker Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Defoerze, Laura

APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C288
CURRENT APPLICATION NUMBER: US/10/176,921
CURRENT FILING DATE: 2002-06-20
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 59
LENGTH: 2550
TYPE: DNA
ORGANISM: Homo Sapien
US-10-176-921-59

Query Match 86.3%; Score 16.4; DB 9; Length 2550;
Best Local Similarity 94.4%; Pred. No. 80;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 GCTCTGATTCCTGCTCTG 19
Db 438 GCTCTGATTACTGCTCTG 455

RESULT 9
US-10-137-865-59
Sequence 59, Application US/10137865
Publication No. US20030032155A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C154
CURRENT APPLICATION NUMBER: US/10/137,865
CURRENT FILING DATE: 2002-05-03
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 59
LENGTH: 2550
TYPE: DNA
ORGANISM: Homo Sapien
US-10-137-865-59

Query Match 86.3%; Score 16.4; DB 9; Length 2550;
Best Local Similarity 94.4%; Pred. No. 80;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 GCTCTGATTCCTGCTCTG 19
Db 438 GCTCTGATTACTGCTCTG 455

RESULT 10
US-10-140-474-59
Sequence 59, Application US/10140474
Publication No. US20030032156A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C162
CURRENT APPLICATION NUMBER: US/10/140,474
CURRENT FILING DATE: 2002-05-06
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 59
LENGTH: 2550
TYPE: DNA
ORGANISM: Homo Sapien
US-10-140-474-59

Query Match 86.3%; Score 16.4; DB 9; Length 2550;
Best Local Similarity 94.4%; Pred. No. 80;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GCTCTGATTCCTGCTCTG 19
Db 438 GCTCTGATTACTGCTCTG 455

RESULT 11
US-10-142-431-59
Sequence 59, Application US/10142431
Publication No. US20030036179A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin

```

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P330R1C251
; CURRENT APPLICATION NUMBER: US/10/142,431
; CURRENT FILING DATE: 2002-05-10
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 59
; LENGTH: 2550
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-142-431-59

Query Match      86.3%; Score 16.4; DB 9; Length 2550;
Best Local Similarity 94.4%; Pred. No. 80;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 GCTCGATTCTGCTCTG 19
         |||||
Db      438 GCTCGATTACTGCTCTG 455

RESULT 12
US-10-143-114-59
; Sequence 59, Application US/10143114
; Publication No. US20030036180A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Zhuang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C211
; CURRENT APPLICATION NUMBER: US/10/143,114
; CURRENT FILING DATE: 2002-05-09
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 59
; LENGTH: 2550
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-143-114-59

Query Match      86.3%; Score 16.4; DB 9; Length 2550;
Best Local Similarity 94.4%; Pred. No. 80;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 GCTCGATTCTGCTCTG 19
         |||||
Db      438 GCTCGATTACTGCTCTG 455

RESULT 13
US-10-140-002-59
; Sequence 59, Application US/10140002
; Publication No. US20030037623A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
```

```

; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Zhuang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C244
; CURRENT APPLICATION NUMBER: US/10/142,419
; CURRENT FILING DATE: 2002-05-10
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 59
; LENGTH: 2550
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-142-419-59

Query Match      86.3%; Score 16.4; DB 9; Length 2550;
Best Local Similarity 94.4%; Pred. No. 80;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 GCTCGATTCTGCTCTG 19
         |||||
Db      438 GCTCGATTACTGCTCTG 455

RESULT 14
US-10-142-419-59
; Sequence 59, Application US/10142419
; Publication No. US20030044945A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Zhuang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C244
; CURRENT APPLICATION NUMBER: US/10/142,419
; CURRENT FILING DATE: 2002-05-10
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 59
; LENGTH: 2550
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-142-419-59

Query Match      86.3%; Score 16.4; DB 9; Length 2550;
Best Local Similarity 94.4%; Pred. No. 80;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

Best Local Similarity 94.4%; Pred. No. 80;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GCTCTGATTCTGCTCTG 19
|||||
Db 438 GCTCTGATTACTGCTCTG 455

RESULT 15

US-10-123-262-59
; Sequence 59, Application US/10123262
; Publication No. US20030049816A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Inc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary B.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C38
; CURRENT APPLICATION NUMBER: US/10/123,262
; CURRENT FILING DATE: 2002-04-15
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 59
; LENGTH: 2550
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-123-262-59

Query Match 86.3%; Score 16.4; DB 9; Length 2550;
Best Local Similarity 94.4%; Pred. No. 80;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GCTCTGATTCTGCTCTG 19
|||||
Db 438 GCTCTGATTACTGCTCTG 455

Search completed: June 21, 2003, 21:12:24
Job time : 122 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: June 21, 2003, 19:32:08 ; Search time 63 Seconds
(without alignments)
92.490 Million cell updates/sec

Title: US-09-581-500B-14

Perfect score: 19

Sequence: 1 cgctcgtatcctgctgctg 19

Scoring table: IDENTITY NTC

Gapop 10.0, Gapext 1.0

Searched: 441362 segs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seg length: 0

Maximum DB seg length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
C 1	15.8	83.2 1200	1 US-08-096-623A-19	Sequence 19, Appl
C 2	15.8	83.2 71989	4 US-09-443-501A-2	Sequence 2, Appl1
C 3	15.8	83.2 246240	2 US-08-724-394A-20	Sequence 20, Appl
C 4	15.8	83.2 246240	2 US-08-724-394A-21	Sequence 21, Appl
C 5	15.8	83.2 246240	2 US-08-724-394A-22	Sequence 22, Appl
C 6	15.4	81.1 1093	1 US-08-592-126-95	Sequence 95, Appl
C 7	15.4	81.1 36519	3 US-08-923-137-2	Sequence 2, Appl1
C 8	15	78.9 1800	1 US-08-139-937-11	Sequence 11, Appl
C 9	15	78.9 1800	5 PCT-US93-11310-11	Sequence 11, Appl
C 10	14.8	77.9 2015	1 US-08-129-129-7	Sequence 7, Appl1
C 11	14.8	77.9 5232	4 US-08-972-927-1	Sequence 1, Appl1
C 12	14.8	77.9 7326	1 US-08-194-468-1	Sequence 1, Appl1
C 13	14.8	77.9 7326	4 US-09-514-247A-7	Sequence 7, Appl1
C 14	14.8	77.9 7344	4 US-08-961-739-1	Sequence 1, Appl1
C 15	14.8	77.9 7431	4 US-09-306-988-2	Sequence 2, Appl1
C 16	14.8	77.9 8147	4 US-09-514-247A-9	Sequence 9, Appl1
C 17	14.8	77.9 16063	4 US-08-801-052-3	Sequence 3, Appl1
C 18	14.8	77.9 45716	4 US-08-965-048-5	Sequence 5, Appl1
C 19	14.8	77.9 45989	4 US-08-965-048-6	Sequence 6, Appl1
C 20	14.4	75.8 1272	3 US-09-058-389A-7	Sequence 7, Appl1
C 21	14.4	75.8 1272	4 US-09-611-781-7	Sequence 7, Appl1
C 22	14.4	75.8 1847	3 US-09-058-389A-9	Sequence 9, Appl1
C 23	14.4	75.8 1847	4 US-09-611-781-9	Sequence 9, Appl1
C 24	14.4	75.8 2522	3 US-09-058-389A-1	Sequence 1, Appl1
C 25	14.4	75.8 2522	4 US-09-611-781-1	Sequence 1, Appl1
C 26	14.4	75.8 3489	2 US-08-728-323A-1	Sequence 1, Appl1
C 27	14.4	75.8 3489	4 US-09-298-568-1	Sequence 1, Appl1

28	14.4	75.8	3970	1 US-07-925-695-3	Sequence 3, Appl1
C 29	14.4	75.8	6354	4 US-09-058-389A-5	Sequence 5, Appl1
C 30	14.4	75.8	6354	3 US-09-611-781-5	Sequence 19, Appl
C 31	14.4	75.8	6803	3 US-08-665-259-19	Sequence 1, Appl1
C 32	14.4	75.8	6803	3 US-08-762-500-19	Sequence 1, Appl1
C 33	14.4	75.8	9412	2 US-08-955-138-1	Sequence 2, Appl1
C 34	14.4	75.8	9589	1 US-07-925-695-1	Sequence 2, Appl1
C 35	14.4	75.8	9589	1 US-07-925-695-2	Sequence 2, Appl1
C 36	14.4	75.8	32207	4 US-08-770-379-20	Sequence 20, Appl
C 37	14.4	75.8	32207	4 US-08-757-669A-20	Sequence 20, Appl
C 38	14.4	75.8	32207	4 US-09-230-371A-20	Sequence 11, Appl
C 39	14.2	74.7	36	3 US-08-808-881-11	Sequence 11, Appl
C 40	14.2	74.7	36	3 US-09-017-631-11	Sequence 11, Appl
C 41	14.2	74.7	66	1 US-08-383-743A-11	Sequence 11, Appl
C 42	14.2	74.7	66	5 PCT-US93-07116-11	Sequence 11, Appl
C 43	14.2	74.7	200	4 US-09-255-368-3	Sequence 3, Appl1
C 44	14.2	74.7	529	1 US-08-253-155A-55	Sequence 5, Appl1
C 45	14.2	74.7	654	1 US-08-390-858B-8	Sequence 8, Appl1

ALIGNMENTS

RESULT 1
US-08-096-623A-19/c
Sequence 19, Application US/08096623A
Patent No. 5684238
GENERAL INFORMATION:
APPLICANT: Ausich, Rodney L.
APPLICANT: Brinkhaus, Friedhelm L.
APPLICANT: Mukharji, Indrani
APPLICANT: Profit, John H.
APPLICANT: Yarger, James G.
APPLICANT: Yen, Hui-Che B.
TITLE OF INVENTION: Biosynthesis of Zeaxanthin and
TITLE OF INVENTION: Glycosylated Zeaxanthin in Genetically Engineered Hosts
NUMBER OF SEQUENCES: 104
CORRESPONDENCE ADDRESS:
ADDRESSEE: Welsh & Katz, Ltd.
STREET: 120 S. Riverside Plaza, 22nd Floor
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/096,623A
FILING DATE: 22-JUL-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/805,061
FILING DATE: 09-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/662,921
FILING DATE: 28-FEB-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/562,674
FILING DATE: 03-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/525,551
FILING DATE: 18-MAY-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/487,613
FILING DATE: 02-MAR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Gamson, Edward P.
REGISTRATION NUMBER: 29,381
REFERENCE/DOCKET NUMBER: AMO-006.1
TELECOMMUNICATION INFORMATION:


```

? TELEPHONE: (312) 655-1500
? TELEFAX: (312) 655-1501
? INFORMATION FOR SEQ ID NO: 19:
? SEQUENCE CHARACTERISTICS:
?   LENGTH: 1200 base pairs
?   TYPE: nucleic acid
?   STRANDEDNESS: single
?   TOPOLOGY: linear
?   MOLECULE TYPE: DNA (genomic)
?   POSITION IN GENOME:
?   MAP POSITION: 1 to 1200
?   UNITS: bp
? US-08-096-623A-19

Query Match      83.2%; Score 15.8; DB 1; Length 1200;
Best Local Similarity 89.5%; Pred. No. 64;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 CGCTCGATTCTGCTCTG 19
        |||||
Db      146 CGCTCGATTCTGCTCTG 128

RESULT 2
US-09-443-501A-2/c
; Sequence 2, Application US/09443501A
; Patent No. 6303342
; GENERAL INFORMATION:
;   APPLICANT: Kosan Biosciences, Inc.
;   APPLICANT: Julien, Bryan
;   APPLICANT: Katz, Leonard
;   APPLICANT: Khosla, Chaitan
;   APPLICANT: Tang, Li
;   APPLICANT: Ziermann, Rainer
;   TITLE OF INVENTION: Recombinant Methods and Materials for Producing
;   TITLE OF INVENTION: Epochilone and Epochilone Derivatives
;   FILE REFERENCE: 30062-20031.00
;   CURRENT APPLICATION NUMBER: US/09/443,501A
;   CURRENT FILING DATE: 1999-11-19
;   PRIOR APPLICATION NUMBER: US 60/130,560
;   PRIOR FILING DATE: 1999-04-22
;   PRIOR APPLICATION NUMBER: US 60/122,620
;   PRIOR FILING DATE: 1999-03-03
;   PRIOR APPLICATION NUMBER: US 60/119,386
;   PRIOR FILING DATE: 1999-02-10
;   PRIOR APPLICATION NUMBER: US 60/109,401
;   PRIOR FILING DATE: 1998-11-20
;   NUMBER OF SEQ ID NOS: 22
;   SOFTWARE: FastSeq for Windows Version 4.0
;   SEQ ID NO 2
;   LENGTH: 71989
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic construct
; US-09-443-501A-2

Query Match      83.2%; Score 15.8; DB 4; Length 71989;
Best Local Similarity 89.5%; Pred. No. 1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```

? APPLICANT: Lauer, Peter M.
? APPLICANT: Ruddy, David A.
? APPLICANT: Thomas, Winston
? APPLICANT: Tsuchihashi, Zenta
? APPLICANT: Wolff, Roger K.
? TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e1
? TITLE OF INVENTION: Sequences and Antibodies Thereo
? NUMBER OF SEQUENCES: 31
? CORRESPONDENCE ADDRESS:
?   ADDRESS: TOWNSEND and TOWNSEND and CREW LLP
?   STREET: Two Embarcadero Center, 8th Floor
?   CITY: San Francisco
?   STATE: CA
?   COUNTRY: USA
?   ZIP: 94111-3834
? COMPUTER READABLE FORM:
?   MEDIUM TYPE: Floppy disk
?   OPERATING SYSTEM: PC-DOS/MS-DOS
?   SOFTWARE: PatentIn Release #1.0, Version #1.30
?   CURRENT APPLICATION DATA:
?   APPLICATION NUMBER: US/08/724,394A
?   FILING DATE: 01-OCT-1996
?   CLASSIFICATION: 536
?   ATTORNEY/AGENT INFORMATION:
?   NAME: Fitts, Renee A.
?   REGISTRATION NUMBER: 35,136
?   REFERENCE/DOCKET NUMBER: 017957-000100
?   TELECOMMUNICATION INFORMATION:
?   TELEPHONE: 415-576-0200
?   INFORMATION FOR SEQ ID NO: 20:
?   SEQUENCE CHARACTERISTICS:
?   LENGTH: 246240 base pairs
?   TYPE: nucleic acid
?   STRANDEDNESS: not relevant
?   TOPOLOGY: not relevant
?   MOLECULE TYPE: cDNA
?   FEATURE:
?   NAME/KEY: misc feature
?   LOCATION: 1..246240
?   OTHER INFORMATION: /note="HLA-H. CONTIG"
? US-08-724-394A-20

Query Match      83.2%; Score 15.8; DB 2; Length 246240;
Best Local Similarity 89.5%; Pred. No. 1.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 CGCTCGATTCTGCTCTG 19
        |||||
Db      161737 CGCTCGATTCTGCTCTG 161755

RESULT 4
US-08-724-394A-21
; Sequence 21, Application US/08724394A
; Patent No. 5872237
; GENERAL INFORMATION:
;   APPLICANT: Feder, John N.
;   APPLICANT: Krommal, Gregory S.
;   APPLICANT: Lauer, Peter M.
;   APPLICANT: Ruddy, David A.
;   APPLICANT: Thomas, Winston
;   APPLICANT: Tsuchihashi, Zenta
;   APPLICANT: Wolff, Roger K.
;   TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e1
;   TITLE OF INVENTION: Sequences and Antibodies Thereo
;   NUMBER OF SEQUENCES: 31
;   CORRESPONDENCE ADDRESS:
;   ADDRESS: TOWNSEND and TOWNSEND and CREW LLP
;   STREET: Two Embarcadero Center, 8th Floor
;   CITY: San Francisco
;   STATE: CA
```

COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/724,394A
FILING DATE: 01-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Filtz, Renee A.
REGISTRATION NUMBER: 35,136
REFERENCE/DOCKET NUMBER: 017957-000100
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0200
TELEFAX: 415-576-0300
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 246240 base pairs
TYPE: nucleic acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..246240
OTHER INFORMATION: /note="HLA-H.CONTIG"
US-08-724-394A-21

Query Match 83.2%; Score 15.8; DB 2; Length 246240;
Best Local Similarity 89.5%; Pred. No. 1.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGCTGTGATTCCTGCTCTG 19
|||||

Db 161737 CGCTGTGATTCCTGCTCTG 161755

RESULT 5
US-08-724-394A-22
Sequence 22, Application US/08724394A
Patent No. 5872237
GENERAL INFORMATION:
APPLICANT: Feder, John N.
APPLICANT: Krommal, Gregory S.
APPLICANT: Laner, Peter M.
APPLICANT: Ruddy, David A.
APPLICANT: Thomas, Winston
APPLICANT: Teuchhashi, Zenta
APPLICANT: Wolff, Roger K.
TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e1
TITLE OF INVENTION: Sequences and Antibodies Thereto
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESSES:
ADDRESS: TOWNSEND and TOWNSEND and CREW LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/724,394A
FILING DATE: 01-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Filtz, Renee A.

REGISTRATION NUMBER: 35,136
REFERENCE/DOCKET NUMBER: 017957-000100
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0200
TELEFAX: 415-576-0300
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 246240 base pairs
TYPE: nucleic acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..246240
OTHER INFORMATION: /note="HLA-H.CONTIG"
US-08-724-394A-22

Query Match 83.2%; Score 15.8; DB 2; Length 246240;
Best Local Similarity 89.5%; Pred. No. 1.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGCTGTGATTCCTGCTCTG 19
|||||

Db 161737 CGCTGTGATTCCTGCTCTG 161755

RESULT 6
US-08-592-126-95/c
Sequence 95, Application US/08592126
Patent No. 5821091
GENERAL INFORMATION:
APPLICANT: Gregory Doljanov
TITLE OF INVENTION: Transcripts Encoding Immunomodulatory
TITLE OF INVENTION: Polypeptides
NUMBER OF SEQUENCES: 151
CORRESPONDENCE ADDRESSES:
ADDRESS: Dellinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/592,126
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 4600-0111
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 95:
SEQUENCE CHARACTERISTICS:
LENGTH: 1093 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: G164con.seq
US-08-592-126-95

Query Match 81.1%; Score 15.4; DB 1; Length 1093;

Best Local Similarity 94.1%; Pred. No. 97;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 2 GCTCGATTCCTGCTCT 18
Db 988 GCTCGATTCCTGCTCT 972

RESULT 7
US-08-923-137-2/C
Sequence 2, Application US/08923137
Patent No. 6083716
GENERAL INFORMATION:
APPLICANT: Wilson, James M.
APPLICANT: Parina, Steven F.
APPLICANT: Fisher, Krishna J.
TITLE OF INVENTION: Chimpanzee Adenovirus Vectors
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Cntr., P.O. Box 457
CITY: Spring House
STATE: Pennsylvania
COUNTRY: United States of America
ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/923,137
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/024,700
FILING DATE: 06-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: GNVN.021CIPUSA
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 36519 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: CDNA
US-08-923-137-2
Query Match 81.1%; Score 15.4; DB 3; Length 36519;
Best Local Similarity 94.1%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 3 CTCTGATTCCTGCTCTG 19
Db 23653 CTCTGATTCCTGCTCTG 23637
RESULT 8
US-08-139-937-11/C
Sequence 11, Application US/08139937
Patent No. 5821070
GENERAL INFORMATION:
APPLICANT: LEE, WEN-HWA
APPLICANT: SHAN, BEI
TITLE OF INVENTION: CELLULAR GENES ENCODING
TITLE OF INVENTION: RETINOBLASTOMA-ASSOCIATED PROTEINS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:

ADDRESSEE: CAMPBELL AND FLORES
STREET: 4370 LA JOLLA VILLAGE DRIVE
CITY: SAN DIEGO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/139,937
FILING DATE: 20-OCT-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/979,156
FILING DATE: 20-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: CAMPBELL, CATRYN
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-CJ 9370
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-535-9001
TELEFAX: 619-535-8949
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1800 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-139-937-11

Query Match 78.5%; Score 15; DB 1; Length 1800;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 2 GCTCGATTCCTGCTCT 16
Db 1698 GCTCGATTCCTGCTCT 1684

RESULT 9
PCT-US93-11310-11/C
Sequence 11, Application PC/TUS9311310
GENERAL INFORMATION:
APPLICANT: BOARD OF REGENTS OF THE UNIVERSITY OF TEXAS SYSTEM
TITLE OF INVENTION: CELLULAR GENES ENCODING
TITLE OF INVENTION: RETINOBLASTOMA-ASSOCIATED PROTEINS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: CAMPBELL AND FLORES
STREET: 4370 LA JOLLA VILLAGE DRIVE
CITY: SAN DIEGO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/11310
FILING DATE: 19-NOV-1993
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: CAMPBELL, CATRYN
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-P-CJ 9790
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-535-9001

TELEX: 619-535-8949
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1800 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
PCT-US93-11310-11

Query Match 78.9%; Score 15; DB 5; Length 1800;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XY 2 GCTCTGATTCCTCTCT 16
|||||
Db 1698 GCTCTGATTCCTCTCT 1684

RESULT 10
US-08-129-129-7
; Sequence 7, Application US/08129129
; Patent No. 5767363
; GENERAL INFORMATION:
; APPLICANT: DE SILVA, Jacqueline
; APPLICANT: SAFFORD, Richard
; APPLICANT: HUGHES, Stephen Glynn
; TITLE OF INVENTION: PLANT PROMOTER INVOLVED IN
; TITLE OF INVENTION: CONTROLLING LIPID BIOSYNTHESIS IN SEEDS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DAREY & CUSHMAN, L.L.P.
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/129,129
; FILING DATE: 18-JAN-1994
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 91303098.7
; FILING DATE: 09-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB92/00627
; FILING DATE: 08-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, Paul N.
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: 203424/T7016(C)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627CUSH
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2015 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 994..1938
US-08-129-129-7

Query Match 77.9%; Score 14.8; DB 1; Length 2015;
Best Local Similarity 88.9%; Pred. No. 2e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

XY 2 GCTCTGATTCCTGCTCTG 19
|||||
Db 1468 GCTCTGATTCCTGCTCTG 1485

RESULT 11
US-08-972-927-1
; Sequence 1, Application US/08972927
; Patent No. 6166290
; GENERAL INFORMATION:
; APPLICANT: Rea, Philip A
; APPLICANT: Lu, Yu-Ping
; APPLICANT: Li, Ze-Sheng
; TITLE OF INVENTION: GLUTATHIONE-S-CONJUGATE TRANSPORT IN
; TITLE OF INVENTION: PLANTS
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PANITCH SCHWARZ JACOBS & NADIEL, P.C.
; STREET: One Commerce Square, 2005 Market Street, 22nd
; STREET: Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: US
; ZIP: 19103-7086
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/972,927
; FILING DATE: 18-NOV-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/031,040
; FILING DATE: 18-NOV-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/061,328
; FILING DATE: 08-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Doyle Leary Ph.D., Kathryn
; REGISTRATION NUMBER: 36,317
; REFERENCE/DOCKET NUMBER: 9596-1202
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-965-1284
; TELEFAX: 215-567-2991
; TELEX: 831-494
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5232 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-972-927-1

Query Match 77.9%; Score 14.8; DB 4; Length 5232;
Best Local Similarity 88.9%; Pred. No. 2.3e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

XY 2 GCTCTGATTCCTGCTCTG 19
|||||
Db 1259 GCTCTGATTCCTGCTCTG 1276

RESULT 12
US-08-194-468-1/c
; Sequence 1, Application US/08194468

Patent No. 5750336
GENERAL INFORMATION:
APPLICANT: Montminy, Marc R.
TITLE OF INVENTION: ASSAYS FOR THE IDENTIFICATION OF
TITLE OF INVENTION: COMPOUNDS WHICH INHIBIT ACTIVATION OF CAMP AND NITROGEN
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
STREET: 444 South Flower Street, Suite 2000
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/194,468
FILING DATE: 10-FEB-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Reiter, Stephen E.
REGISTRATION NUMBER: 31,192
REFERENCE/DOCKET NUMBER: P41 9672
TELEPHONE: (619)-546-4737
TELEFAX: (619)-546-9392
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 7326 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHEICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..7323
US-08-194-468-1
Query Match 77.9%; Score 14.8; DB 1; Length 7326;
Best Local Similarity 88.9%; Pred. No. 2.3e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 GCTGTGATTCCTGCTCTG 19
Db 3136 GCTGTGATTCCTGCTCTG 3119
RESULT 13
US-09-514-247A-7/c
Sequence 7, Application US/09514247A
Patent No. 6365361
GENERAL INFORMATION:
APPLICANT: TANABE SEIYAKU CO. LTD.
APPLICANT: TANIGUCHI, Tomoyasu
APPLICANT: MIZUKAMI, Junko
TITLE OF INVENTION: METHOD FOR IDENTIFYING OR SCREENING AGONIST AND ANTAGONIST TO PPA
FILE REFERENCE: TANIGUCHI=6
CURRENT APPLICATION NUMBER: US/09/514,247A
CURRENT FILING DATE: 2000-02-28
PRIOR APPLICATION NUMBER: PCT/JP98/03734
PRIOR FILING DATE: 1998-08-24
PRIOR APPLICATION NUMBER: JP231084/1997
PRIOR FILING DATE: 1997-08-27
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn version 3.0
SEQ ID NO 7
LENGTH: 7326

TYPE: DNA
ORGANISM: mouse
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(7326)
OTHER INFORMATION: n at position 1131 is unknown.
US-09-514-247A-7
Query Match 77.9%; Score 14.8; DB 4; Length 7326;
Best Local Similarity 88.9%; Pred. No. 2.3e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 GCTGTGATTCCTGCTCTG 19
Db 3136 GCTGTGATTCCTGCTCTG 3119
RESULT 14
US-08-961-739-1/c
Sequence 1, Application US/08961739A
Patent No. 6063583
GENERAL INFORMATION:
APPLICANT: Montminy, Marc R.
TITLE OF INVENTION: Methods for Treating Diabetes Mellitus
FILE REFERENCE: SALK1650-1
CURRENT APPLICATION NUMBER: US/08/961,739A
CURRENT FILING DATE: 1997-10-31
EARLIER APPLICATION NUMBER: US 194,468
EARLIER FILING DATE: 1994-02-10
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 7326
TYPE: DNA
ORGANISM: Mus
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(7326)
NAME/KEY: misc feature
LOCATION: (1)..(7326)
OTHER INFORMATION: n = A,T,C or G
US-08-961-739-1
Query Match 77.9%; Score 14.8; DB 3; Length 7344;
Best Local Similarity 88.9%; Pred. No. 2.3e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 GCTGTGATTCCTGCTCTG 19
Db 3136 GCTGTGATTCCTGCTCTG 3119
RESULT 15
US-09-306-998-2
Sequence 2, Application US/09306998
Patent No. 6291173
GENERAL INFORMATION:
APPLICANT: Bartel, Paul L.
APPLICANT: Tavtigian, Sean V.
TITLE OF INVENTION: NMSC2 - An MMAC1 Interacting Protein
FILE REFERENCE: NMSC2
CURRENT APPLICATION NUMBER: US/09/306,998
CURRENT FILING DATE: 1999-05-07
EARLIER APPLICATION NUMBER: US 60/084,740
EARLIER FILING DATE: 1998-05-08
NUMBER OF SEQ ID NOS: 72
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 7431
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS

LOCATION: (57) .. (6167)
US-09-306-998-2

```
Query Match      77.9%; Score 14.8; DB 4; Length 743;
Best Local Similarity 88.9%; Pwd. No. 2.4e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

QY 2 GCTCTGATTCCTGCTCTG 19
Db 303 GCTGTGATTCCTACTCTG 320

Search completed: June 21, 2003, 20:21:43
Job time : 65 secs

PT Detecting nucleotide triplet repeats in human chromosome 18q
 SS Claim 29, Fig 15b, 87pp; English.
 XX
 CC The present invention describes detecting nucleotide triplet repeats in
 CC a region of human chromosome 18q disposed between polymorphic markers
 CC D18S68 and D18S979 to identify a human gene associated with a mood
 CC disorder or related disorder. AAX88542 to AAX88705 represents human
 CC chromosome 18q YAC clones and primers corresponding to them, used in the
 CC exemplification of the present invention. YAC clones comprising a
 CC portion of the region of human chromosome 18q between markers D18S68 and
 CC D18S979 are used to identify at least one human gene associated with a
 CC mood disorder or related disorder. The mood disorder or related
 CC disorder, is chosen from the Diagnostic and Statistical Manual of Mental
 CC Disorders, version 4 (DSM-IV) taxonomy. This includes mood disorders
 CC (296.XX, 300.4, 311, 301, 13, 295.70), schizophrenia and related
 CC disorders (295, 297.1, 298.9, 297.3, 298.9), anxiety disorders (300.XX,
 CC 309.81, 308.3), adjustment disorders (309.XX) and personality disorders
 CC (codes 301.XX). Probes derived from genes associated with the mood
 CC disorder or related disorder can be used to detect pathological
 CC mutations or genetic variations in patients. The methods, probes and
 CC antibodies can be used to determine the susceptibility of an individual
 CC to a mood disorder or related disorder. The nucleic acids and proteins
 CC of the human gene can be used to treat mood disorders and related
 CC disorders.

SQ Sequence 19 BP; 1 A; 7 C; 4 G; 7 T; 0 other;

Query Match 100.0%; Score 19; DB 20; Length 19;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCTGATTCCTGCTCTG 19
 Db 1 CGCTGATTCCTGCTCTG 19

RESULT 2
 AAX88553/c
 ID AAX88553 standard; DNA; 656 BP.

AC AAX88553;
 DT 10-SEP-1999 (first entry)
 DE Human chromosome 18q YAC clone nucleotide sequence #12.
 XX
 KW Human chromosome 18q; mood disorder; polymorphic marker; detection;
 KW identification; trinucleotide repeat expansion; schizophrenia;
 KW anxiety disorder; adjustment disorder; personality disorder;
 KW nucleotide triplet repeat; ss.
 XX
 OS Homo sapiens.
 OS Synthetic.
 PN WO9932643-A2.
 PD 01-JUL-1999.
 PF 17-DEC-1998; 98WO-EP08543.
 PR 18-DEC-1997; 97GB-0026804.
 PA (VIAA-) VLAMS INTERNUNIVERSITAIR INST BIOTECHNOG.
 PI Del-Favero J, Raeymaekers P, Van Broeckhoven C;
 DR WPI; 1999-418934/35.
 PT Detecting nucleotide triplet repeats in human chromosome 18q
 XX
 PS Claim 21, Fig 15a; 87pp; English.
 XX

CC The present invention describes detecting nucleotide triplet repeats in
 CC a region of human chromosome 18q disposed between polymorphic markers
 CC D18S68 and D18S979 to identify a human gene associated with a mood
 CC disorder or related disorder. AAX88542 to AAX88705 represents human
 CC chromosome 18q YAC clones and primers corresponding to them, used in the
 CC exemplification of the present invention. YAC clones comprising a
 CC portion of the region of human chromosome 18q between markers D18S68 and
 CC D18S979 are used to identify at least one human gene associated with a
 CC mood disorder or related disorder. The mood disorder or related
 CC disorder, is chosen from the Diagnostic and Statistical Manual of Mental
 CC Disorders, version 4 (DSM-IV) taxonomy. This includes mood disorders
 CC (296.XX, 300.4, 311, 301, 13, 295.70), schizophrenia and related
 CC disorders (295, 297.1, 298.9, 297.3, 298.9), anxiety disorders (300.XX,
 CC 309.81, 308.3), adjustment disorders (309.XX) and personality disorders
 CC (codes 301.XX). Probes derived from genes associated with the mood
 CC disorder or related disorder can be used to detect pathological
 CC mutations or genetic variations in patients. The methods, probes and
 CC antibodies can be used to determine the susceptibility of an individual
 CC to a mood disorder or related disorder. The nucleic acids and proteins
 CC of the human gene can be used to treat mood disorders and related
 CC disorders.

SQ Sequence 656 BP; 183 A; 165 C; 208 G; 96 T; 4 other;

Query Match 100.0%; Score 19; DB 20; Length 656;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCTGATTCCTGCTCTG 19
 Db 533 CGCTGATTCCTGCTCTG 515

RESULT 3
 ABR3567
 ID ABR3567 standard; cDNA; 112460 BP.

AC ABR3567;
 DT 14-AUG-2002 (first entry)
 DE Human cDNA differentially expressed in granulocytic cells #138.
 XX
 KW Human; ss; granulocytic cell; DNA chip; bacterial infection;
 KW viral infection; parasitic infection; protozoal infection;
 KW fungal infection; streptococcal inflammatory disease; psoriasis;
 KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
 KW cardiac reperfusion injury; renal reperfusion injury; ARDS;
 KW adult respiratory distress syndrome; inflammatory bowel disease;
 KW Crohn's disease; ulcerative colitis; periodontal disease;
 KW granulocyte activation; chronic inflammation; allergy.
 XX
 OS Homo sapiens.
 PN WO200228999-A2.
 PD 11-APR-2002.
 PF 03-OCT-2001; 2001WO-US30821.
 PR 03-OCT-2000; 2000US-237189P.
 PA (GENE-) GENA LOGIC INC.
 PI Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
 DR WPI; 2002-435328/46.
 PT Detecting granulocyte activation by detecting differential expression
 PT of genes associated with granulocyte activation, which serves as
 PT diagnostic markers that is useful for monitoring disease states and
 PT drug toxicity -
 XX

PS Claim 1; SEQ ID No 138; 114pp; English.

XX The invention relates to detecting (M1) granulocyte (GC) activation

CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by

CC DNA chip analysis as given in the specification, and comparing

CC the expression level to an expression level in an unactivated

CC GC, where differential expression of Gs is indicative of GCA.

CC Also included are modulating (M2) GA by contacting GC with an agent

CC that alters the expression of at least one gene in Gs; (2) screening (M3)

CC for an agent capable of modulating GCA or an inflammation (especially

CC chronic) in a tissue, an allergic response in a subject, exposure of a

CC subject to a pathogen or sterile inflammatory disease using the

CC gene expression profile; (3) detecting (M4) an inflammation (especially

CC chronic) in a tissue, an allergic response in a subject, exposure of a

CC subject to a pathogen or sterile inflammatory disease, by detecting the

CC level of expression in a sample of the tissue of gene(s) from Gs, where

CC the level of expression of the gene is indicative of inflammation;

CC (4) treating (M5) an inflammation (especially chronic) or in a tissue,

CC or allergic response in a subject, exposure of a subject to a pathogen

CC or sterile inflammatory disease, by contacting a tissue having

CC inflammation with an agent that modulates the expression of gene(s)

CC from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for

CC modulating GA; M3 is useful for screening an agent capable of modulating

CC GCA preferably in an inflammation in a tissue; M4 is useful for

CC detecting an inflammation (especially chronic) in a tissue, an allergic

CC response in a subject, exposure of a subject to a pathogen or sterile

CC inflammatory disease (e.g. psoriasis, rheumatoid arthritis,

CC glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal

CC reperfusion injury, ARDS, adult respiratory distress syndrome,

CC inflammatory bowel disease, Crohn's disease, ulcerative colitis,

CC periodontal disease, also bacterial infection, viral infection, and MS is

CC parasitic infection, protozoal infection, fungal infection and MS is

CC useful for treating one of the above conditions. The present

CC sequence represents a gene differentially expressed in granulocytes.

CC Note: The sequence data for this patent did not form part

CC of the printed specification, but was obtained in electronic

CC format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences.

XX

XX Sequence 112460 BP; 24087 A; 29523 C; 31203 G; 27647 T; 0 other;

XX

XX Query Match 89.5%; Score 17; DB 24; Length 112460;

XX Best Local Similarity 100.0%; Pred. No. 2.4e+02;

XX Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX

XX 3 CTCGTGATTCCTGCTCTG 19

XX |||||

XX 88792 CTCGTATTCCTGCTCTG 88808

XX

XX RESULT 4

XX AAF93539

XX ID AAF93539 standard; cDNA; 502 BP.

XX AC AAF93539;

XX DT 21-MAY-2001 (first entry)

XX XX

XX Lung carcinoma cDNA encoding SRT protein SEQ ID 360.

XX DE

XX Human; SRT; gene therapy; gene mapping; tissue typing; ss.

XX KM

XX Homo sapiens.

XX OS

XX WO200107611-A2.

XX PN

XX 01-FEB-2001.

XX PD

XX 21-JUL-2000; 2000WO-US20006.

XX PF

XX 26-JUL-1999; 99US-0145701.

XX PR

XX (SETH) GENENTECH INC.

XX PA

XX Baker KP, Goddard A, Wood WI;

XX WPI; 2001-112729/12.

XX

XX New isolated nucleic acid molecule encoding a SRT polypeptide is useful

XX PT for production of recombinant SRT polypeptides, gene mapping,

XX PT diagnosing genetic disorders and for gene therapy -

XX PS

XX Claim 2; Fig 360; 663pp; English.

XX

XX Sequences AAF93180 - AAF93743 represent polynucleotide sequences encoding

XX human SRT proteins. The cDNA sequences are isolated from various

XX different human tissue cDNA libraries. The invention relates to a method

XX for detecting cDNA encoding an SRT protein, a vector containing cDNA

XX encoding SRT, a host cell transformed with the vector, an isolated SRT

XX polypeptide, and an antibody which binds to SRT. The polynucleotide

XX sequence can be used in gene therapy and is useful in the recombinant

XX production of SRT polypeptides, as a hybridization probe to screen

XX libraries to isolate cDNAs with sequence identity to SRT polypeptides, to

XX map the gene encoding the SRT polypeptides and analysing genetic

XX disorders, tissue typing and disease detection. The SRT

XX polynucleotide sequences can be used in polymerase chain reaction,

XX screening for new therapeutic molecules and generation of antisense RNA

XX and DNA.

XX

XX Sequence 502 BP; 94 A; 149 C; 119 G; 131 T; 9 other;

XX

XX Query Match 86.3%; Score 16.4; DB 22; Length 502;

XX Best Local Similarity 94.4%; Pred. No. 3.2e+02;

XX Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX

XX 2 GCTCTGATTCCTGCTCTG 19

XX |||||

XX 400 GCTCTGATTCCTGCTCTG 417

XX

XX RESULT 5

XX AAF93542

XX ID AAF93542 standard; cDNA; 511 BP.

XX AC AAF93542;

XX DT 21-MAY-2001 (first entry)

XX XX

XX Lung carcinoma cDNA encoding SRT protein SEQ ID 363.

XX DE

XX Human; SRT; gene therapy; gene mapping; tissue typing; ss.

XX KM

XX Homo sapiens.

XX OS

XX WO200107611-A2.

XX PN

XX 01-FEB-2001.

XX PD

XX 21-JUL-2000; 2000WO-US20006.

XX PF

XX 26-JUL-1999; 99US-0145701.

XX PR

XX (SETH) GENENTECH INC.

XX PA

XX Baker KP, Goddard A, Wood WI;

XX WPI; 2001-112729/12.

XX

XX New isolated nucleic acid molecule encoding a SRT polypeptide is useful

XX PT for production of recombinant SRT polypeptides, gene mapping,

XX PT diagnosing genetic disorders and for gene therapy -

XX PS

XX Claim 2; Fig 363; 663pp; English.

XX

XX Sequences AAF93180 - AAF93743 represent polynucleotide sequences encoding

XX human SRT proteins. The cDNA sequences are isolated from various

CC and cancer and for promoting wound healing. The sequences are used to
CC induce the proliferation of neural cells and regeneration of nerve and
CC brain tissue, and are useful for the treatment of central and peripheral
CC nervous system diseases and neuropathies, such as Alzheimer's disease,
CC Parkinson's disease, Huntington's disease and amyotrophic lateral
CC sclerosis. The sequences are involved in chemotactic or chemokinetic
CC activity, regulation of hematopoiesis, treatment of myeloid or lymphoid
CC cell disorders and platelet disorders such as thrombocytopenia,
CC regeneration of bone, cartilage, tendon, ligament and/or nerve tissue
CC growth, tissue repair, healing of burns, incisions, ulcers, treatment of
CC osteoporosis, osteoarthritis, bone degenerative disorders and periodontal
CC disease. The sequences of the invention are also useful for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues, immune deficiencies and disorders
CC including severe combined immunodeficiency (SCID), bacterial or fungal
CC infections, autoimmune disorders e.g. multiple sclerosis and myasthenia
CC gravis, allergic conditions such as asthma, thrombolytic or thrombotic
CC and coagulation disorders. Sequences ABR94890-ABR94982 represent human
CC novel polynucleotides of the invention.

XX SQ Sequence 1769 BP; 389 A; 480 C; 483 G; 417 T; 0 other;

Query Match 86.3%; Score 16.4; DB 24; Length 1769;

Best Local Similarity 94.4%; Pred. No. 3.5e+02; Mismatches 1; Indels 0; Gaps 0;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GCTCGATTCCTGCTCTG 19

Db 424 GCTCGATTCCTGCTCTG 441

RESULT 8

AAH13643/c
ID AAH13643 standard; cDNA; 1873 BP.

XX AC AAH13643;

XX DT 26-JUN-2001 (first entry)

XX DE Human cDNA sequence SEQ ID NO:10485.

XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX OS Homo sapiens.

XX PN EP1074617-A2.

XX PD 07-FEB-2001.

XX PF 28-JUL-2000; 2000EP-0116126.

XX PR 29-JUL-1999; 99JP-0248036.

XX PR 27-AUG-1999; 99JP-0300253.

XX PR 11-JAN-2000; 2000JP-0118776.

XX PR 02-MAY-2000; 2000JP-0183767.

XX PR 09-JUN-2000; 2000JP-0241899.

XX PA (HELI-) HELIX RES INST.

XX PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX DR WPI; 2001-318749/34.

XX PT Primer sets for synthesizing polynucleotides, particularly the 5602

XX PT full-length cDNAs defined in the specification, and for the detection

XX PT and/or diagnosis of the abnormality of the proteins encoded by the

XX PT full-length cDNAs -

XX PS Claim 8; SEQ ID 10485; 2537PP + CD ROM; English.

XX CC The present invention describes primer sets for synthesizing 5602

XX CC full-length cDNAs defined in the specification. Where a primer set

CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH1628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.

XX SQ Sequence 1873 BP; 549 A; 419 C; 526 G; 379 T; 0 other;

Query Match 86.3%; Score 16.4; DB 22; Length 1873;

Best Local Similarity 94.4%; Pred. No. 3.5e+02; Mismatches 1; Indels 0; Gaps 0;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GCTCGATTCCTGCTCTG 19

Db 914 GCTTGAATCCTGCTCTG 897

RESULT 9

AAH16697
ID AAH16697 standard; cDNA; 2252 BP.

XX AC AAH16697;

XX DT 26-JUN-2001 (first entry)

XX DE Human cDNA sequence SEQ ID NO:15859.

XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX OS Homo sapiens.

XX PN EP1074617-A2.

XX PD 07-FEB-2001.

XX PF 28-JUL-2000; 2000EP-0116126.

XX PR 29-JUL-1999; 99JP-0248036.

XX PR 27-AUG-1999; 99JP-0300253.

XX PR 11-JAN-2000; 2000JP-0118776.

XX PR 02-MAY-2000; 2000JP-0183767.

XX PR 09-JUN-2000; 2000JP-0241899.

XX PA (HELI-) HELIX RES INST.

XX PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX DR WPI; 2001-318749/34.

XX PT Primer sets for synthesizing polynucleotides, particularly the 5602

XX PT full-length cDNAs defined in the specification, and for the detection

XX PT and/or diagnosis of the abnormality of the proteins encoded by the

XX PT full-length cDNAs -

XX PS Claim 8; SEQ ID 15859; 2537PP + CD ROM; English.

XX CC The present invention describes primer sets for synthesizing 5602

CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dt primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination
 CC of the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialising methods. AAH03166 to AAH1628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.

XX
 SQ Sequence 2252 BP; 560 A; 567 C; 608 G; 517 T; 0 other;

Query Match 86.3%; Score 16.4; DB 22; Length 2252;
 Best Local Similarity 94.4%; Pred. No. 3.5e+02;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GCTCTGATTCCTGCTCTG 19
 |||||
 Db 141 GCTCTGATTCCTGCTCTG 158

RESULT 10
 AAS21273
 ID AAS21273 standard; cDNA; 2550 BP.
 XX
 AC AAS21273;
 XX
 DT 24-OCT-2001 (first entry)
 XX
 XX Human cDNA sequence encoding for PRO1779 polypeptide.
 XX
 KW Human secretory and transmembrane; PRO; mammalian; cancer; lung;
 KW breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;
 KW cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;
 KW adipocyte; A-peptide; factor VIIA; gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 PN WC0200140466-A2.
 XX
 PD 07-JUN-2001.
 XX
 XX 01-DEC-2000; 2000WO-US32678.
 PF
 XX 01-DEC-1999; 99WO-US28301.
 PR 01-DEC-1999; 99WO-US28634.
 PR 02-DEC-1999; 99WO-US28551.
 PR 02-DEC-1999; 99WO-US28564.
 PR 02-DEC-1999; 99WO-US28565.
 PR 09-DEC-1999; 99US-0170262.
 PR 16-DEC-1999; 99WO-US30095.
 PR 20-DEC-1999; 99WO-US30911.
 PR 20-DEC-1999; 99WO-US30999.
 PR 30-DEC-1999; 99WO-US31243.
 PR 06-JAN-2000; 2000WO-US00277.
 PR 06-JAN-2000; 2000WO-US00376.
 PR 11-FEB-2000; 2000WO-US04365.
 PR 18-FEB-2000; 2000WO-US04341.
 PR 18-FEB-2000; 2000WO-US04342.
 PR 22-FEB-2000; 2000WO-US04414.
 PR 24-FEB-2000; 2000WO-US04914.

PR 24-FEB-2000; 2000WO-US05004.
 PR 01-MAR-2000; 2000WO-US05601.
 PR 20-MAR-2000; 2000WO-US07377.
 PR 21-MAR-2000; 2000WO-US07532.
 PR 30-MAR-2000; 2000WO-US08439.
 PR 17-MAY-2000; 2000WO-US13705.
 PR 22-MAY-2000; 2000WO-US14042.
 PR 30-MAY-2000; 2000WO-US14941.
 PR 02-JUN-2000; 2000WO-US15264.
 PR 10-NOV-2000; 2000WO-US30873.
 XX
 PA (GENTECH) GENENTECH INC.
 XX
 PI Baker KP, Beresini M, DeGeorge L, Filvaroff E, Gao W,
 PI Gerlitsen ME, Goddard A, Godowski PJ, Gunney AL, Sherwood S;
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WT, Zhang Z;
 XX
 DR WPI; 2001-408281/43.
 DR P-PSDB; AAU12201.
 XX
 PT Isolated, secretory and transmembrane PRO polypeptide used to detect
 PT other PRO polypeptides, link bioactive molecules to cells expressing
 PT PRO polypeptides, and detect the presence of mammalian tumours e.g.
 PT lung, breast, prostate, cervical
 XX
 PS Claim 3; Fig 59; 81pp; English.
 XX
 CC AAS21244-AAS21518 encode for novel human secretory and transmembrane
 CC PRO polypeptides. The PRO polypeptides are useful to detect other
 CC PRO polypeptides, to link bioactive molecules to cells expressing
 CC PRO polypeptides, to modulate biological activities of cells expressing
 CC PRO polypeptides, and to detect the presence of mammalian lung, colon,
 CC breast, prostate, rectal, cervical or liver tumours by comparing PRO
 CC polypeptide expression in a cell sample to that in a control sample.
 CC Some of the 275 sequences are also useful to stimulate the release of
 CC tumour necrosis factor-alpha (TNF-alpha) from human blood, the
 CC proliferation or differentiation of chondrocytes, the proliferation or
 CC gene expression in pericyte cells, the release of proteoglycans from
 CC cartilage, the proliferation of inner ear utricular supporting cells or
 CC of T-lymphocytes, the release of a cytokine from peripheral blood
 CC monocytes (PBMCs), or the proliferation of endothelial cells. Some of
 CC the PRO polypeptides may modulate glucose or free fatty acid uptake by
 CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide
 CC to factor VIIA. The PRO polypeptides can be used in assays to identify
 CC molecules involved in binding interactions. The polynucleotides encoding
 CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,
 CC transgenic or knock out animals and can be used in gene therapy.

XX
 SQ Sequence 2550 BP; 610 A; 666 C; 683 G; 591 T; 0 other;

Query Match 86.3%; Score 16.4; DB 22; Length 2550;
 Best Local Similarity 94.4%; Pred. No. 3.5e+02;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GCTCTGATTCCTGCTCTG 19
 |||||
 Db 438 GCTCTGATTCCTGCTCTG 455

RESULT 11
 AAA80606
 ID AAA80606 standard; cDNA; 2609 BP.
 XX
 AC AAA80606;
 XX
 DT 21-NOV-2000 (first entry)
 XX
 XX Human secreted protein gene #1.
 DE
 XX Secreted protein; immunosuppressant; anti-inflammatory; antiarthritic;
 KW antihemetic, dermatological; antiproliferative; antitumorleucotic;
 KW anticancer; vulnery; antiviral; antibacterial; antifungal;
 KW immune disorder; Addison's disease; rheumatoid arthritis; dermatitis;

KW	multiple sclerosis; inflammatory disorder; inflammatory bowel disease;
KM	Crohn's disease; nephritis; hyperproliferative disorder;
KV	cardiovascular disorder; coronary arteriosclerosis; myocarditis; cancer;
KX	melanoma; lymphoma; wound healing; human; ss.
XX	
OS	Homo sapiens.
XX	
FN	WC000029435-A1.
PD	
XX	25-MAY-2000.
XX	
PF	27-OCT-1999; 99NO-US25031.
XX	
PR	28-OCT-1998; 98US-0105971.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
PI	
PI	M J, Ruben SM, Olsen HS, Young PE, Kenny JT, Moore PA, Wel Y,
XX	Greene JM;
DR	WPI; 2000-387742/33.
XX	P-PDB; AAB25576.
PT	
PT	Isolated nucleic acid molecules encoding human secreted proteins are
PT	used for the prevention, amelioration and treatment of autoimmune,
PT	inflammatory, hyperproliferative and cardiovascular disorders, cancer,
PT	wounds, and infectious diseases -
XX	
PS	Claim 1, Figure 1A-C; 803pp; English.
XX	
CC	The present invention relates to 12 secreted human proteins and the
CC	nucleotide sequences encoding them. The polynucleotide sequences given
CC	in AA08606-AA0623 encode the 12 secreted protein sequences given in
CC	AAB25576-BB25933. The human secreted proteins have various activities
CC	dependent on the tissues in which they are expressed. Examples of the
CC	activities of the proteins include: immunosuppressant;
CC	anti-inflammatory; antirheumatic; antineoplastic, dermatological;
CC	antiproliferative; antiarteriosclerotic; anticancer; vulnery;
CC	antiviral; antibacterial; and antifungal activity. The proteins,
CC	polypeptides, agonists and antagonists may be used to treat prevent
CC	and/or diagnose various disease, disorders and conditions examples of
CC	which include: immune disorders e.g. Addison's disease, rheumatoid
CC	arthritis, dermatitis, and multiple sclerosis; inflammatory disorders
CC	e.g. inflammatory bowel disease, Crohn's disease and nephritis;
CC	hyperproliferative disorders such as paraneoplasmas and purpura;
CC	cardiovascular disorders e.g. coronary arteriosclerosis and myocarditis;
CC	cancer e.g. melanoma and lymphoma. The proteins and polynucleotide
CC	sequences may also be used in wound healing and the treatment of
CC	infectious diseases. The human secreted protein gene #1 and protein
CC	sequences are represented in sequences AA0606 and AAB25576. Also given
CC	is an alternative protein product of human secreted protein gene #1 in
CC	AAB25594.
XX	
SQ	Sequence 2609 BP; 682 A; 658 C; 660 G; 589 T; 0 other;
XX	
Query Match	86.3%; Score 16.4; DB 21; Length 2609;
Best Local Similarity	94.4%; Pred. No. 3.5e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0	
OY	2 GCCTGATTCCTGCTTG 19
Db	411 GCCTGATTACTGCTCG 428
RESULT_12	
AAD05402/c	
ID	AAD05402 standard; CDNA; 3185 BP.
AC	
XX	AAD05402;
XX	
DT	17-JUL-2001 (first entry)
DE	Human secreted protein-encoding gene 14 cDNA clone HPJKBK11, SEQ ID NO:24.

KW	Human; secreted protein; proliferative disorder; cancer; tumour;
KX	foetal abnormality; developmental abnormality; haematopoietic disorder;
KM	immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KM	inflammation; allergy; neurological disorder; Alzheimer's disease;
KM	Parkinson's disease; cognitive disorder; schizophrenia; asthma;
KM	skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
KM	cardiovascular disorder; angiotensin disorder; kidney disorder;
KM	gastrointestinal disorder; pregnancy-related disorder;
KM	endocrine disorder; infection; wound healing; vulnerability;
KM	cell culture; chemotaxis; food additive; chromosome 9;
KW	binding partner identification; gene therapy; ss.
XN	
OS	Homo sapiens.
XX	
XX	Key Location/Qualifiers
FT	CDS 169..459
FT	/tag= a
FT	/product= "Human secreted protein"
FT	169..252
FT	/tag= b
FT	253..456
FT	/tag= c
FT	/product= "Mature human secreted protein"
PN	
XX	WO200134623-A1.
PD	
XX	17-MAY-2001.
PP	01-NOV-2000; 2000NC-US30037.
XX	
PR	05-NOV-1999; 99US-0163577.
XX	
PR	30-JUN-2000; 2000US-0215137.
XX	
PA	(HDA-) HUMAN GENOME SCT INC.
PI	
DR	Ruben SM, Komatsoulis GA, Moore PA;
XX	
DR	WPI; 2001-316490/33.
PT	P-RSDB; AAE01559.
XX	
XX	Nucleic acids encoding 29 human secreted polypeptides, useful for
XX	preventing, diagnosing and/or treating e.g. cancers, Parkinson's
XX	disease and diabetic retinopathy -
XX	
XX	Claim 1: Page 423-424; 535pp; English.
CC	AAD05389-AA05473 represent cDNAs corresponding to 29 human secreted
CC	protein genes, and AAE01546-AAE01630 represent the proteins they encode.
CC	AAE01631-AAE01660 represent human secreted protein fragments or variants
CC	The secreted proteins and their genes are useful for preventing,
CC	treating or ameliorating medical conditions, e.g., by protein or gene
CC	therapy. Pathological conditions can be diagnosed by determining the
CC	amount of the new protein in a sample or by determining the presence of
CC	mutations in the new genes. Specific uses are described for each of the
CC	29 genes, based on the tissues in which they are most highly expressed,
CC	and include developing products for the diagnosis or treatment of
CC	proliferative disorders, cancer, tumours, foetal and developmental
CC	abnormalities, haematopoietic disorders, diseases of the immune system,
CC	AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
CC	allergies, neurological disorders (e.g., Alzheimer's disease,
CC	Parkinson's disease), cognitive disorders, schizophrenia, asthma,
CC	skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
CC	cardiovascular disorders, angiogenic disorders, kidney disorders,
CC	gastrointestinal disorders, pregnancy-related disorders, endocrine
CC	disorders, and infections. The proteins can also be used to aid wound
CC	healing and epithelial cell proliferation, to prevent skin aging due to
CC	sunburn, to maintain organs before transplantation, for supporting cell
CC	culture of primary tissues, to regenerate tissues, to identify their
CC	cognate ligands or binding partners, and in chemotaxis, and can be used
CC	as a food additive or preservative to modify storage properties.
CC	Antibodies specific for a protein of the invention can be used in
CC	alleviating symptoms associated with the disorders mentioned above, and

CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
 CC immunosorbent assay (ELISA). The present sequence represents a human
 CC secreted protein-encoding cDNA of the invention.

XX Sequence 3185 BP, 772 A, 808 C, 890 G, 715 T, 0 other;

Query Match 86.3%; Score 16.4; DB 22; Length 3185;
 Best Local Similarity 94.4%; Pred. No. 3.6e+02;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GCTCTGATTCCTCTCTG 19
 Db 1515 GCTCAGATCTCTCTCTG 1498

RESULT 13
 ABQ8177/c
 ID ABQ8177 standard; cDNA; 125439 BP.

XX ABQ8177;

DT 18-SEP-2002 (first entry)

DE Human osteoblast differentiation related cDNA SEQ ID NO 84.

XX Human; osteoblast; stem cell differentiation; bone tissue deposition;

KW osteoporosis; osteopathic; ss.

XX Homo sapiens.

XX WO200250301-A2.

PD 27-JUN-2002.

PF 18-DEC-2001; 2001WO-US48276.

PR 18-DEC-2000; 2000US-255882P.

PR 24-APR-2001; 2001US-285691P.

XX (GENE-) GENE LOGIC INC.

PA (PROCT) PROCTER & GAMBLE CO.

PI J1 D, Axelrod DW, Cook JS, Jaiswal N, Binstein R, Houghton A;

PI Meritz L;

PT MPI; 2002-557663/59.

DR Use of genes and their expression profiles associated with osteoblast

PT differentiation for screening modulators bone formation, for diagnosing

PT or treating e.g. osteoporosis, or as markers for the differentiation

PT process

XX Claim 1, SEQ ID NO 84; 78bp + Sequence listing; English.

XX The invention relates to genes and their expression profiles are used

CC for:

CC (a) screening modulators of precursor stem cell differentiation into

CC osteoblasts, or bone tissue deposition; abnormal rate of

CC (b) diagnosing abnormal deposition of bone tissue, abnormal rate of

CC osteoblast formation or osteoporosis; or

CC (c) treating or monitoring treatment of the conditions cited in (b), or

CC monitoring the progression of bone tissue deposition.

CC Specific conditions include postmenopausal osteoporosis, glucocorticoid

CC osteoporosis or male osteoporosis, osteopenia, osteodystrophy,

CC drug-induced abnormalities in bone formation or bone loss, conditions

CC that involve altered bone metabolism (e.g. idiopathic juvenile

CC osteoporosis), skeletal disease linked to breast cancer, mastocytosis,

CC Panconi syndrome or fibrous dysplasia. The present sequence is that of an

CC osteoblast differentiation associated cDNA marker of the invention.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 125439 BP; 32323 A; 32592 C; 30004 G; 30520 T; 0 other;

Query Match 86.3%; Score 16.4; DB 24; Length 125439;
 Best Local Similarity 94.4%; Pred. No. 4.4e+02;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GCTCTGATTCCTCTCTG 19
 Db 49494 GCTCTGATTCCTCTCTG 49477

RESULT 14
 ABN45912/c
 ID ABN45912 standard; DNA; 60 BP.

XX ABN45912;

DT 15-JUL-2002 (first entry)

DE Human spliced transcript detection oligonucleotide SEQ ID NO:18660.

XX Human; mouse; rat; splice transcript; detection; RNA transcript;

KW splice variant; transcriptome; oligonucleotide library; ss.

XX Homo sapiens.

XX WO200210449-A2.

PD 07-FEB-2002.

PF 20-JUL-2001; 2001WO-IB01903.

PR 28-JUL-2000; 2000US-221607P.

PR 02-MAY-2001; 2001US-287724P.

XX (COMP-) COMPUGEN INC.

PA Shoshan A, Wasserman A, Mintz B, Mintz L, Faigler S;

PI MPI; 2002-257383/30.

DR New oligonucleotide libraries comprising oligonucleotides which

PT selectively hybridize to mRNAs transcribed from a transcription unit of

PT a genome, useful for detecting tissue-, pathology-, and

PT developmental-specific genes

XX Example 1; SEQ ID 18660; 47bp; English.

XX The present invention describes oligonucleotide libraries for detecting

CC messenger RNAs that populate a (sub-)transcriptome, where the

CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple

CC transcription units that populate a genome. The library comprises

CC several oligonucleotides, each capable of hybridizing selectively to a

CC set of messenger RNAs transcribed from a given transcription unit of

CC the genome, which encodes one or more messenger RNA splice variants.

CC The oligonucleotide libraries are useful for detecting mRNAs from a

CC biological sample, in expression profiling studies, in qualitatively or

CC quantitatively characterizing the corresponding transcriptome, and in

CC detecting RNA transcripts and splice variants of human or animal

CC transcriptomes. The libraries may also be used as specialised mini

CC libraries to detect transcripts of a sub-transcriptome under a

CC particular biological or pathological state, and so allowing the

CC detection of tissue- or pathology-specific genes such as those genes

CC only expressed in specific tissue under a specific pathological

CC condition; to detect developmental specific genes; and to detect RNA

CC transcripts and splice variants of a transcriptome of a patient suffering

CC from a particular disorder. ABN27253 to ABN59589 represent

CC oligonucleotide sequences from rats, humans and mice, which are used in

CC the exemplification of the present invention.

CC N.B. The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

Sequence 60 BP; 9 A; 17 C; 18 G; 16 T; 0 other;

Query Match 83.2%; Score 15.8; DB 24; Length 60;
Best Local Similarity 89.5%; Pred. No. 5.3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 CGCTCGATTCCTGCTG 19
50 CGCTCGATTCCTGCTG 32

RESULT 15

AAT24088/c
ID AAT24088 standard; cDNA to mRNA; 151 BP.

AAT24088;

06-SEP-1996 (first entry)

Human gene signature HMG806076.

Gene signature; messenger RNA; mRNA; relative abundance; frequency;
human; cloning; mapping; non-biased library; diagnosis; detection;
cell typing; abnormal cell function; ss.

Homo sapiens.

MO9514772-A1.

01-JUN-1995.

11-NOV-1994; 94WO-JP01916.

12-NOV-1993; 93JP-0355504.

(MATS/) MATSUBARA K.

(OKUB/) OKUBO K.

Matsubara K, Okubo K;

WPI; 1995-206931/27.

Identifying gene signatures in 3'-directed human cDNA library - e.g.
for diagnosis of abnormal cell function, by preparing cDNA that
reflects relative abundance of corresp. mRNA in specific human
tissues

Claim 1; Page 1527; 2245bp; Japanese.

A single-stranded DNA (or its complementary strand or the corresp.
double-stranded DNA) which comprises one of the 7837 "GS" sequences
given in AAT19001-T26837 and which is able to hybridise to part of
human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
sequences were obtained from 3'-directed cDNA libraries prepared
from various human tissues; synthesis of cDNA was initiated from the
3'-end of mRNA by using poly(U) as the sole primer. Since the 3'-
untranslated sequence is unique to a particular mRNA species, almost
all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
is constructed so as to reflect accurately the relative abundance of
different mRNAs in the particular tissue from which it was derived.
The appearance frequency of a given GS in a cDNA library can be
determined (esp. using primers and probes derived from the GS
sequences) as a means of diagnosing abnormal cell function or for
recognising different cell types.

Sequence 151 BP; 49 A; 33 C; 39 G; 29 T; 1 other;

Query Match 83.2%; Score 15.8; DB 16; Length 151;
Best Local Similarity 89.5%; Pred. No. 5.6e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 CGCTCGATTCCTGCTG 19
151 CGCTCGATTCCTGCTG 151

DB 25 CGCTCGATTCCTGCTG 7

Search completed: June 21, 2003, 19:39:34
Job time : 218 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2003, 18:31:18 ; Search time 991 Seconds
(without alignments)
557.975 Million cell updates/sec

Title: US-09-581-500B-14

Perfect score: 19

Sequence: 1 cgcctcgtatccctgcctcg 19

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmb1.*
1: gb_da.*
2: gb_hcg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*
15: em_da.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vl.*
30: em_hcg_hum.*
31: em_hcg_inv.*
32: em_hcg_other.*
33: em_hcg_mus.*
34: em_hcg_pln.*
35: em_hcg_rnd.*
36: em_hcg_mam.*
37: em_hcg_wrt.*
38: em_sy.*
39: em_hngo_hum.*
40: em_hngo_mus.*
41: em_hngo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19	100.0	19	AX021008	AX021008 Sequence
2	19	100.0	656	AX021006	AX021006 Sequence
3	19	100.0	186351	AP001503	AP001503 Homo sapi
4	19	100.0	181385	AC009802	AC009802 Homo sapi
5	19	100.0	191793	AP001897	AP001897 Homo sapi
6	19	100.0	198291	AC067875	AC067875 Homo sapi
7	19	100.0	201734	AC090224	AC090224 Homo sapi
8	17.4	91.6	2121	AF140243	AF140243 Xenopus 1
9	17.4	91.6	151130	AC127432	AC127432 Mus muscu
10	17.4	91.6	151395	AC128860	AC128860 Rattus no
11	17.4	91.6	168369	AC119463	AC119463 Rattus no
12	17.4	91.6	176054	AC022377	AC022377 Homo sapi
13	17.4	91.6	178548	AC018831	AC018831 Homo sapi
14	17.4	91.6	189040	AC125748	AC125748 Rattus no
15	17.4	91.6	192180	AL845471	AL845471 Mus muscu
16	17.4	91.6	196044	AC018505	AC018505 Homo sapi
17	17.4	91.6	198597	AC026218	AC026218 Homo sapi
18	17.4	91.6	198808	AC129207	AC129207 Mus muscu
19	17.4	91.6	202844	AC090956	AC090956 Homo sapi
20	17.4	91.6	203026	AL772271	AL772271 Mus muscu
21	17.4	91.6	204412	AC021446	AC021446 Mus muscu
22	17.4	91.6	205925	AC009711	AC009711 Homo sapi
23	17.4	91.6	209876	AC011599	AC011599 Homo sapi
24	17.4	91.6	229402	AC069271	AC069271 Homo sapi
25	17.4	91.6	256657	AC129177	AC129177 Mus muscu
26	17.4	89.5	12693	U58729	U58729 Caenorhabdi
27	17	89.5	112460	HS1104E15	AL022312 Human DNA
28	17	89.5	159659	AC116868	AC116868 Mus muscu
29	17	89.5	169193	AC068523	AC068523 Homo sapi
30	17	89.5	176526	AP005230	AP005230 Homo sapi
31	17	89.5	196500	AC068734	AC068734 Homo sapi
32	17	89.5	208844	AC118200	AC118200 Mus muscu
33	17	89.5	262387	AL670959	AL670959 Mouse DNA
34	16.4	86.3	305	AF027365	AF027365 Rattus no
35	16.4	86.3	502	AX079616	AX079616 Sequence
36	16.4	86.3	511	AX079619	AX079619 Sequence
37	16.4	86.3	540	CNS01B18	AL14264 Botrytis
38	16.4	86.3	901	CNS06EK4	AL195210 T3 end of
39	16.4	86.3	1268	BC015201	BC015201 Homo sapi
40	16.4	86.3	1873	AK000870	AK000870 Homo sapi
41	16.4	86.3	2252	AK023902	AK023902 Homo sapi
42	16.4	86.3	2343	BC034347	BC034347 Mus muscu
43	16.4	86.3	2514	BC016552	BC016552 Homo sapi
44	16.4	86.3	2550	AX463926	AX463926 Sequence
45	16.4	86.3	2638	BC025476	BC025476 Mus muscu

ALIGNMENTS

RESULT 1
LOCUS AX021008
DEFINITION Sequence 14 from Patent W09932643.
ACCESSION AX021008
VERSION AX021008.1 GI:10044671
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 19)
AUTHORS Del-Favero,U., Raeymaekers,P. and Van Broeckhoven,C.
TITLE Mood disorder gene
JOURNAL Patent: WO 9932643-A 14 01-JUL-1999;

FEATURES
SOURCE
1.19
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT
1 a 7 c 4 g 7 t

ORIGIN

Query Match 100.0%; Score 19; DB 6; Length 19;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY
1 CGCTCGATTCCTGCTCG 19
1 CGCTCGATTCCTGCTCG 19

Db
1 CGCTCGATTCCTGCTCG 19

RESULT 2
AX021006/c
LOCUS
DEFINITION
Accession
AX021006
Version
AX021006.1 GI:10044669
Keywords
SOURCE
Homo sapiens
human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 656)
Del-Pavero,J., Raeymaekers,P. and Van Broeckhoven,C.
Mood disorder gene
Patent: WO 9932643-A.12 01-JUL-1999;
BROECKHOVEN CHRISTINE VAN (BE); DEL-PAVERO JURGEN (BE); RAEYMAEKERS
PETER (BE); VLAAMS INTERUNIV INST BIOTECH (BE)
Location/Qualifiers
1..656
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT
183 a 165 c 208 g 96 t 4 others

ORIGIN

Query Match 100.0%; Score 19; DB 6; Length 656;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY
1 CGCTCGATTCCTGCTCG 19
533 CGCTCGATTCCTGCTCG 515

Db
533 CGCTCGATTCCTGCTCG 515

RESULT 3
AP001503/c
LOCUS
DEFINITION
Accession
AP001503
Version
AP001503.2 GI:8117361
Keywords
HTG; HTGS PHASE1; HTGS DRAFT.
Homo sapiens DNA, clone:RP11-879N20.
Homo sapiens
human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 186351)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Torokiy,Y., Matanabe,H. and Sakaki,Y.
Homo sapiens 186,351 genomic DNA of 18q22
Published Only in Database (2000)
2 (bases 1 to 186351)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Torokiy,Y., Matanabe,H. and Sakaki,Y.
Direct Submission
Submitted (17-MAR-2000) Masahira Hattori, The Institute of Physical

COMMENT

and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
Japan (E-mail:hattori@gs.c.riken.go.jp)
URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,
Fax:81-42-778-9924)
On May 30, 2000 this sequence version replaced gi:7288199.
----- Genome Center -----
Center: RIKEN Genomic Sciences Center (GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gs.c.riken.go.jp
----- Project Information -----
Center project name: HumDrafit18
Center clone name: RP11-879N20
----- Summary Statistics -----
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 162162 bases at least Q40
Consensus quality: 175389 bases at least Q30
Consensus quality: 181326 bases at least Q20
Insert size: 184151, sum-of-contrigs
Quality coverage: 4.54x in Q20 bases; sum-of-contrigs

NOTE: This is a 'working draft' sequence. It currently consists of
23 contrigs. The true order of the pieces is not known and their
order in this sequence record is arbitrary. Gaps between the
contrigs are represented as runs N, but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved.
1 42519 contrig of 42519 bp in length
42620 59852 contrig of 17233 bp in length
59953 74496 contrig of 14544 bp in length
74597 87440 contrig of 12844 bp in length
87541 99577 contrig of 12037 bp in length
99678 110659 contrig of 10982 bp in length
110760 115514 contrig of 4654 bp in length
115514 124990 contrig of 9477 bp in length
125091 135744 contrig of 10654 bp in length
135845 145972 contrig of 10128 bp in length
146073 153170 contrig of 7088 bp in length
153271 156398 contrig of 3128 bp in length
156499 161724 contrig of 5226 bp in length
161825 165688 contrig of 3864 bp in length
165789 169698 contrig of 3910 bp in length
169799 173554 contrig of 37354 bp in length
173655 175764 contrig of 2110 bp in length
175865 177303 contrig of 1439 bp in length
177404 179097 contrig of 1694 bp in length
179198 181636 contrig of 2439 bp in length
181737 183644 contrig of 1908 bp in length
183745 184761 contrig of 1017 bp in length
184862 186351 contrig of 1440 bp in length
Sequence updated (26-May-2000).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 23 contrigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contrigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 42519 contrig of 42519 bp in length
42520 42619 contrig of 100 bp
42620 59852 contrig of 17233 bp in length
59853 59952 contrig of 100 bp in length
59953 74496 contrig of 14544 bp in length
74497 74596 contrig of 100 bp
74597 74597 contrig of 100 bp
74597 87440 contrig of 12844 bp in length
87441 87540 contrig of 100 bp
87541 99577 contrig of 12037 bp in length
99578 99677 contrig of 100 bp
99678 99677 gap of 100 bp

TITLE
JOURNAL
COMMENT

McGowan, P., McGurk, A., McKernan, K., McPheters, R., Meldrum, J.,
Menas, L., Mihova, T., Miranda, C., Mlenza, Y., Morrow, J., Naylor, J.,
Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, T.M.,
Peterson, K., Pierre, N., Pisan, C., Pollara, V., Raymond, C.,
Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S.,
Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N.,
Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Tirrell, A.,
Travers, M., Trifilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B.,
Wu, X., Wyman, D., Ye, W.J., Young, G., Zannoun, J., Zimmer, A. and
Zody, M.

Direct Submission

Submitted (17-FEB-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jan 16, 2000 this sequence version replaced gi:5514043.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www.seq.wi.mit.edu>

Contact: sequence.submissions@genome.wi.mit.edu

Project Information

Center project name: L1250

Center clone name: 793_J_2

FEATURES

Source

Location/Qualifiers

1. 191395

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="18"

/map="18"

/clone="RP11-793J2"

/clone_1lb="RP11-793J2"

/clone_1lb="RP11-793J2"

/clone_1lb="RP11-793J2"

/clone_1lb="RP11-793J2"

/clone_1lb="RP11-793J2"

/clone_1lb="RP11-793J2"

/clone_1lb="RP11-793J2"

/clone_1lb="RP11-793J2"

/clone_1lb="RP11-793J2"

/clone_1lb="RP11-793J2"

/clone_1lb="RP11-793J2"

/clone_1lb="RP11-793J2"

/clone_1lb="RP11-793J2"

/clone_1lb="RP11-793J2"

/clone_1lb="RP11-793J2"

/clone_1lb="RP11-793J2"

/clone_1lb="RP11-793J2"

/clone_1lb="RP11-793J2"

/clone_1lb="RP11-793J2"

/clone_1lb="RP11-793J2"

/clone_1lb="RP11-793J2"

/clone_1lb="RP11-793J2"

/clone_1lb="RP11-793J2"

/clone_1lb="RP11-793J2"

/clone_1lb="RP11-793J2"

/clone_1lb="RP11-793J2"

/clone_1lb="RP11-793J2"

/clone_1lb="RP11-793J2"

/clone_1lb="RP11-793J2"

/clone_1lb="RP11-793J2"

/clone_1lb="RP11-793J2"

/clone_1lb="RP11-793J2"

/clone_1lb="RP11-793J2"

/clone_1lb="RP11-793J2"

/clone_1lb="RP11-793J2"

/clone_1lb="RP11-793J2"

/clone_1lb="RP11-793J2"

/clone_1lb="RP11-793J2"

/clone_1lb="RP11-793J2"

/clone_1lb="RP11-793J2"

/clone_1lb="RP11-793J2"

/clone_1lb="RP11-793J2"

/rpt_family="(TATA)n"

repeat_region

5742..5907

/rpt_family="(TATA)n"

repeat_region

complement(5966..6047)

/rpt_family="(TATA)n"

repeat_region

6882..6703

/rpt_family="(TATA)n"

repeat_region

7034..7933

/rpt_family="(TATA)n"

repeat_region

7935..8006

/rpt_family="(TATA)n"

repeat_region

8009..9111

/rpt_family="(TATA)n"

repeat_region

9112..9165

/rpt_family="(TATA)n"

repeat_region

9216..9496

/rpt_family="(TATA)n"

repeat_region

9497..9520

/rpt_family="(TATA)n"

repeat_region

9561..10264

/rpt_family="(TATA)n"

repeat_region

complement(10837..10840)

/rpt_family="(TATA)n"

repeat_region

complement(11156..11515)

/rpt_family="(TATA)n"

repeat_region

11856..11987

/rpt_family="(TATA)n"

repeat_region

12592..12643

/rpt_family="(TATA)n"

repeat_region

13683..13994

/rpt_family="(TATA)n"

repeat_region

complement(13996..14160)

/rpt_family="(TATA)n"

repeat_region

14788..15255

/rpt_family="(TATA)n"

repeat_region

complement(15875..15996)

/rpt_family="(TATA)n"

repeat_region

16101..16135

/rpt_family="(TATA)n"

repeat_region

16136..16304

/rpt_family="(TATA)n"

repeat_region

complement(16572..16647)

/rpt_family="(TATA)n"

repeat_region

complement(16716..17011)

/rpt_family="(TATA)n"

repeat_region

17045..17780

/rpt_family="(TATA)n"

repeat_region

complement(17945..18133)

/rpt_family="(TATA)n"

repeat_region

complement(18222..18755)

/rpt_family="(TATA)n"

repeat_region

18843..19019

/rpt_family="(TATA)n"

repeat_region

19573..19616

/rpt_family="(TATA)n"

repeat_region

20506..20544

/rpt_family="(TATA)n"

repeat_region

21894..22184

/rpt_family="(TATA)n"

repeat_region

complement(22787..22998)

/rpt_family="(TATA)n"

repeat_region

complement(23074..23133)

/rpt_family="(TATA)n"

repeat_region

23184..23540

/rpt_family="(TATA)n"

repeat_region

complement(23675..25225)

/rpt_family="(TATA)n"

repeat_region

25229..25582

/rpt_family="(TATA)n"

repeat_region

25679..26031

/rpt_family="(TATA)n"

repeat_region

complement(26031..26031)

/rpt_family="(TATA)n"

repeat_region

complement(26031..26031)

/rpt_family="(TATA)n"

repeat_region

complement(26031..26031)

/rpt_family="(TATA)n"

repeat_region

complement(26031..26031)

/rpt_family="(TATA)n"

repeat_region

complement(26031..26031)

/rpt_family="(TATA)n"

repeat_region

complement(26031..26031)

/rpt_family="(TATA)n"

repeat_region

complement(26031..26031)

/rpt_family="(TATA)n"

repeat_region

complement(26031..26031)

/rpt_family="(TATA)n"

repeat_region

complement(26031..26031)

/rpt_family="(TATA)n"

repeat_region

complement(26031..26031)

/rpt_family="(TATA)n"

repeat_region

complement(26031..26031)

/rpt_family="(TATA)n"

repeat_region

complement(26031..26031)

/rpt_family="(TATA)n"

repeat_region

complement(26031..26031)

/rpt_family="(TATA)n"

repeat_region

complement(26031..26031)

/rpt_family="(TATA)n"

*	1	10723	contig of	10723	bp in length
*	10824	19450	contig of	8627	bp in length
*	19551	26733	contig of	9189	bp in length
*	28840	37796	contig of	8957	bp in length
*	37897	45643	contig of	7747	bp in length
*	45744	53200	contig of	7457	bp in length
*	53301	59466	contig of	6166	bp in length
*	59567	65022	contig of	5457	bp in length
*	65124	71723	contig of	6602	bp in length
*	71826	77046	contig of	5221	bp in length
*	77147	82059	contig of	4915	bp in length
*	82160	86515	contig of	4355	bp in length
*	86565	90836	contig of	4221	bp in length
*	90937	96257	contig of	4321	bp in length
*	95358	99276	contig of	3919	bp in length
*	99377	103751	contig of	4375	bp in length
*	103852	108407	contig of	4556	bp in length
*	108508	112468	contig of	3841	bp in length
*	112553	116411	contig of	3363	bp in length
*	116512	120690	contig of	4179	bp in length
*	120791	123789	contig of	4959	bp in length
*	125890	129429	contig of	3540	bp in length
*	129550	133875	contig of	3346	bp in length
*	132976	136718	contig of	3743	bp in length
*	136819	140204	contig of	3386	bp in length
*	140305	143212	contig of	2908	bp in length
*	143313	146311	contig of	2999	bp in length
*	146412	148694	contig of	2283	bp in length
*	148795	151551	contig of	2757	bp in length
*	151652	154741	contig of	1820	bp in length
*	153572	155905	contig of	2334	bp in length
*	156006	157930	contig of	1925	bp in length
*	158031	160335	contig of	2305	bp in length
*	160436	161957	contig of	1522	bp in length
*	162058	164115	contig of	2058	bp in length
*	164216	166374	contig of	2159	bp in length
*	166475	168634	contig of	2160	bp in length
*	168735	170796	contig of	2062	bp in length
*	170897	172731	contig of	1835	bp in length
*	172832	174464	contig of	1633	bp in length
*	174565	175680	contig of	1116	bp in length
*	175781	176918	contig of	1138	bp in length
*	177019	178822	contig of	1804	bp in length
*	178923	180077	contig of	1155	bp in length
*	180178	181248	contig of	1071	bp in length
*	181349	182499	contig of	1151	bp in length
*	182600	183806	contig of	207	bp in length
*	182907	183932	contig of	1086	bp in length
*	184093	185198	contig of	1106	bp in length
*	185239	186320	contig of	1022	bp in length
*	186421	187601	contig of	1181	bp in length
*	187702	189059	contig of	1397	bp in length
*	189200	190576	contig of	1378	bp in length
*	190677	191793	contig of	1117	bp in length
Sequence updated (26-May-2000).					
* NOTE: This is a "working draft" sequence. It currently					
* consists of 54 contigs. The true order of the pieces					
* is not known and their order in this sequence record is					
* arbitrary. Gaps between the contigs are represented as					
* runs of N, but the exact sizes of the gaps are unknown.					
* This record will be updated with the finished sequence					
* as soon as it is available and the accession number will					
* be preserved.					
*	1	10723	contig of	10723	bp in length
*	10724	10823	gap of	100	bp
*	10824	19450	contig of	8627	bp in length
*	19451	19550	gap of	100	bp
*	19551	28739	contig of	9189	bp in length
*	28740	28839	gap of	100	bp
*	28840	37796	contig of	8957	bp in length
*	37797	37896	gap of	100	bp
*	37897	45643	contig of	7747	

```

45744 53200: contig of 7457 bp in length
53201 53300: gap of 100 bp
53301 59466: contig of 6166 bp in length
59467 59566: gap of 100 bp
59567 65023: contig of 5457 bp in length
65024 65123: gap of 100 bp
65124 71725: contig of 6602 bp in length
71726 71825: gap of 100 bp
71826 77046: contig of 5221 bp in length
77047 77146: gap of 100 bp
77147 82059: contig of 4913 bp in length
82060 82159: gap of 100 bp
82160 86515: contig of 4356 bp in length
86516 86615: gap of 100 bp
86616 90836: contig of 4221 bp in length
90837 90936: gap of 100 bp
90937 95257: contig of 4321 bp in length
95258 95357: gap of 100 bp
95358 99276: contig of 3919 bp in length
99277 99376: gap of 100 bp
99377 103751: contig of 4375 bp in length
103752 103851: gap of 100 bp
103852 108407: contig of 4556 bp in length
108408 108507: gap of 100 bp
108508 112468: contig of 3961 bp in length
112469 112568: gap of 100 bp
112569 116411: contig of 3843 bp in length
116412 116511: gap of 100 bp
116512 120690: contig of 4179 bp in length
120691 120790: gap of 100 bp
120791 125789: contig of 4999 bp in length
125790 125889: gap of 100 bp
125890 129429: contig of 3540 bp in length
129430 128529: gap of 100 bp
128530 132875: contig of 3346 bp in length
132876 132975: gap of 100 bp
132976 136718: contig of 3743 bp in length
136719 136818: gap of 100 bp
136819 140204: contig of 3386 bp in length
140205 140304: gap of 100 bp
140305 143212: contig of 2908 bp in length
143213 143312: gap of 100 bp
143313 146311: contig of 2999 bp in length
146312 146411: gap of 100 bp
146412 148694: contig of 2283 bp in length
148695 148794: gap of 100 bp
148795 151551: contig of 2757 bp in length
151552 151651: gap of 100 bp
151652 153471: contig of 1820 bp in length
153472 153571: gap of 100 bp
153572 155905: contig of 2334 bp in length
155906 156005: gap of 100 bp
156006 157930: contig of 1929 bp in length
157931 158030: gap of 100 bp
158031 160335: contig of 2305 bp in length
160336 160435: gap of 100 bp
160436 161957: contig of 1522 bp in length
161958 162057: gap of 100 bp
162058 164115: contig of 2058 bp in length
164116 164215: gap of 100 bp
164216 166374: contig of 2159 bp in length
166375 166474: gap of 100 bp
166475 168634: contig of 2160 bp in length
168635 168734: gap of 100 bp
168735 170796: contig of 2062 bp in length
170797 170896: gap of 100 bp
170897 172731: contig of 1835 bp in length
172732 172831: gap of 100 bp
172832 174464: contig of 1633 bp in length
174465 174564: gap of 100 bp
174565 175680: contig of 1116 bp in length
175681 175780: gap of 100 bp
175781 176918: contig of 1138 bp in length

```

```

Query Match 100.0%; Score 19; DB 2; Length 191793;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Ct 1 CGCTGATTCGCTCTG 19
Db 28583 CGCTGATTCGCTCTG 28565

```

```

AC067875 198291 bp DNA linear HMG 24-MAY-2001
Homo sapiens chromosome 18 clone RP11-589B3 map 18, *** SEQUENCING
IN PROGRESS ***, 4 unordered pieces.
AC067875 GI:14193020
HMG: HMG5 PHASE1; HMG5_FULLTOP; HMG5_ACTIVEFIN.
KEYWORDS
Homo sapiens.
ORGANISM
Homo sapiens.

```

```

REFERENCE
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Birren, B., Linton, L., Nusbaum, C. and Lander, E.
1 (bases 1 to 198291)
Homo sapiens chromosome 18, clone RP11-589B3
Unpublished
2 (bases 1 to 198291)

```

```

AUTHORS
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,
Boguski, J., Brown, A., Brown, A., Buckett, G.,
Campione, A., Castle, A., Choehel, Y., Colangelo, M., Collins, S.,
Collins, A., Cooke, P., DeCarliano, K., Dewar, K., Diaz, J.S.,
Dodge, S., Donato, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,
Galagan, J., Gardina, S., Ginde, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karakas, A.,
Klein, J., Labrecque, K., Lamazares, R., Landers, T., Lebeck, J.,
Levine, R., Liu, C., Liu, G., Locke, K., MacDonald, P., Margulis, N.,
McCarthy, M., McGowan, P., McGurk, A., McKernan, K., McPherson, R.,
Meldrum, J., Menus, L., Mihova, T., Miranda, C., Miya, V., Morrow, J.,
Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N.,
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
Strange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Teste, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.-J.,
Young, G., Zaitoun, D., Zimmer, A. and Zody, M.

```

```

TITLE
Direct Submission
JOURNAL
Submitted (27-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 24, 2001 this sequence version replaced gi:13560423.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

```

```

----- Genome Center
Center: Whitehead Institute/MIT Center for Genome Research
Center code: WIRB
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: 18567
Center clone name: 589_E3

```

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence

```

* as soon as it is available and the accession number will
* be preserved.
1 38268: contig of 38268 bp in length
* 38269 38368: gap of 100 bp
* 38369 109704: contig of 71336 bp in length
* 109705 109804: gap of 100 bp in length
* 109805 128704: contig of 18900 bp in length
* 128705 128804: gap of 100 bp
* 128805 198291: contig of 69487 bp in length.
Location/Qualifiers
1. 198291
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="18"
/map="18"
/clone="RP11-589E3"
/clone_1lb="RP11-589E3 Human Male BAC"
BASE COUNT 62849 a 38924 c 38683 g 57534 t 301 others
ORIGIN
Query Match 100.0%; Score 19; DB 2; Length 198291;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 CGCTGTGATTCCTGCTG 19
|||||
Db 182643 CGCTGTGATTCCTGCTG 182661
RESULT 7
AC090224 201734 bp DNA linear HTG 23-MAY-2001
AC090224 Homo sapiens chromosome 18 clone RP11-693A18 map 18, WORKING DRAFT
DEFINITION
SEQUENCE, 8 unordered pieces.
AC090224
AC090224 GI:14190716
HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 201734)
Birren, B., Linton, L., Nusbaum, C. and Lander, E.
Homo sapiens chromosome 18, clone RP11-693A18
REFERENCE
JOURNAL
TITL
AUTHORS
2 (bases 1 to 201734)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,
Barna, N., Basteien, V., Boguski, L., Bouckigalter, B., Brown, A.,
Camata, U., Campopiano, A., Choepel, Y., Colangelo, M., Collins, S.,
Collamore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S.,
Dodg, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J.,
Gardna, S., Ginde, S., Goyette, M., Graham, L., Grand-pierre, N.,
Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
Jones, C., Karatas, A., Laroque, K., Lamares, R., Landers, T.,
Lehoczky, U., Levine, R., Liu, G., Maclean, C., Macdonald, P.,
Margis, N., Mathews, C., McCarthy, M., McEwen, P., McKernan, K.,
McPeters, R., Meltin, J., Menus, L., Mihova, T., Mlang, V.,
Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C. H.,
O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K.,
Phunhng, S., Pierre, N., Pollara, V., Raymond, C., Retta, R.,
Rieback, M., Riley, R., Riese, C., Rogov, P., Roman, T., Rosetti, M.,
Roy, A., Santos, R., Schauer, S., Schumpack, R., Seaman, S., Severy, P.,
Souarez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Struss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodores, J.,
Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Vial, R., Vo, A.,
Wilson, B., Wu, X., Wyman, D., Ye, W., Young, G., Zainoun, J.,
Zemek, L., Zimmer, A. and Zody, M.
Direct Submission
TITLE
JOURNAL
COMMENT
Submitted (17-FEB-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 23, 2001 this sequence version replaced gi:13621272.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIRB
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L1285
Center clone name: 693_A_18
----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 198542 bases at least Q40
Consensus quality: 200270 bases at least Q30
Consensus quality: 200829 bases at least Q20
Insert size: 194000; agarose-fp
Insert size: 201034; sum-of-contigs
Quality coverage: 11.3 in Q20 bases; agarose-fp
Quality coverage: 10.9 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 45183: contig of 45183 bp in length
* 45184 45283: gap of 100 bp
* 45284 46768: contig of 1485 bp in length
* 46769 46868: gap of 100 bp
* 46869 47961: contig of 1093 bp in length
* 47962 48061: gap of 100 bp
* 48062 48643: contig of 1582 bp in length
* 48644 49743: gap of 100 bp
* 49744 51393: contig of 1650 bp in length
* 51394 51493: gap of 100 bp
* 51494 74531: contig of 23038 bp in length
* 74532 74631: gap of 100 bp
* 74632 127710: contig of 53073 bp in length
* 127711 127810: gap of 100 bp
* 127811 201734: contig of 73924 bp in length.
Location/Qualifiers
1. 201734
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="18"
/map="18"
/clone="RP11-693A18"
/clone_1lb="RP11-693A18 Human Male BAC"
misc_feature
1. 45183
/note="assembly_fragment"
clone_end:SP6
vector_side:left"
45284. .46768
/note="assembly_fragment"
46869. .47961
/note="assembly_fragment"
48062. .49643
/note="assembly_fragment"
49744. .51393
/note="assembly_fragment"
51494. .74531
/note="assembly_fragment"
74632. .127710
/note="assembly_fragment"
127811. .201734
/note="assembly_fragment"
clone_end:T7
vector_side:right"
BASE COUNT 61611 a 39444 c 39572 g 60399 t 708 others

ORIGIN

Query Match 100.0%; Score 19; DB 2; Length 201734;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTCGATTCCTGCTG 19
 DB 147537 CGCTCGATTCCTGCTG 147555

RESULT 8
 AF140243/c 2121 bp mRNA linear VRT 07-MAR-2000
 LOCUS AF140243
 DEFINITION Xenopus laevis axin-related protein mRNA, complete cds.
 ACCESSION AF140243
 VERSION AF140243.1 GI:6652990
 KEYWORDS
 SOURCE Xenopus laevis.
 ORGANISM Xenopus laevis.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;

REFERENCE
 1 (bases 1 to 2121)
 Itoh, K., Antipova, A., Ratcliffe, M. J. and Sokol, S.
 Interaction of dishevelled and Xenopus axin-related protein is
 required for wnt signal transduction
 Mol. Cell. Biol. 20 (6), 2228-2238 (2000)

JOURNAL MEDLINE 2013783
 PUBMED 10688669
 REFERENCE 2 (bases 1 to 2121)
 Antipova, A., Itoh, K. and Sokol, S.
 Direct Submission
 Submitted (02-APR-1999) Microbiology and Medical Genetics, Harvard
 Medical School, BIDMC, East Campus, RW663, 330 Brookline Avenue,
 Boston, MA 02215, USA

FEATURES

source

CDS

1..2121
 /organism="Xenopus laevis"
 /db_xref="taxon:8355"
 /issue_type="ovary"
 1..2121
 /note="ARP"
 /codon_start=1
 /product="axin-related protein"
 /protein_id="XARP2574.1"
 /db_xref="GI:6652991"
 /translation="MSSAGVLTICIPDSGPIRRETSLRPPVQGEINPKPKETVDSO
 HIKHEDFNRAECVAHDSRFRGSRINILDDQGAFLFRMYLEGGLDILTEW
 FAKHGFPMADLEPKTSKTKAKAIYMYONSASVGRKLEKTFYQVLECVNKLQNTK
 VPDQOEIOAPMOBAEFTLOSIDCKEYRAGVEDPTESPQGLPLTLEDEFG
 LHHPSGKGKINRAESRIIPRNRSHFRKLCOTQYAPRASTNDEFSDDLTEDM
 SNTDSVDGIPPRSKKOREIHSVSANGKYSLPFPRTRPAEMKSPSPAPAKI
 TIALRKVKQDAEKLBEKLDKLEERLADYDISSSHVTPGALLEDQSTLDD
 HVSRYLKTNPANISPRSQSPVQKGFOPARBEETSAVLTPLSPQEARNSVLTLS
 ELRSVSSQLPRSRKPGCTGPHRPEEGTAAVLTPLSPQEARNSVLTLS
 AKLMKHEHREAVTPCPCEKATRAASOPAHPLDTPMLPLTAPLTDLEBAR
 RLVDKRPYKPKHKRSQVSTLKKEKKAASVPSGSGSTLKLEBQKTKKPSGCG
 QGLATVYFCGERLIPWIRTEPSLTLOEFELLSKGSKNYTFKESHSFPCNAVFG
 EVSEDAVLPLEFKIKCYERAC"

BASE COUNT 637 a 501 c 507 g 476 t
 ORIGIN

Query Match 91.6%; Score 17.4; DB 5; Length 2121;
 Best Local Similarity 94.7%; Pred. No. 1.3e+02;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGCTCGATTCCTGCTG 19
 DB 1490 CGCTCGATTCCTGCTG 1472

RESULT 9

AC127432

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLB

JOURNAL

REFERENCE

AUTHORS

TITLB

JOURNAL

COMMENT

FEATURES

source

BASE COUNT 4308 a 3212 c 32537 g 43383 t
 ORIGIN

Query Match 91.6%; Score 17.4; DB 10; Length 151130;
 Best Local Similarity 94.7%; Pred. No. 1.5e+02;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGCTCGATTCCTGCTG 19
 DB 99788 CGCTCGATTCCTGCTG 99806

RESULT 10

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 151395)
 Wuzny, D. M., Adams, C., Adio-Oduola, B., Ali-oshan, F. R., Allen, C.,
 Alshrook, S. L., Amaral, H. C., Are, J. R., Ayala, M., Banks, T.,
 Barbarta, J., Benton, J., Bimaga, K., Blankenburg, K., Bonnin, D.,
 Bouck, J., Bowls, S., Briefa, M., Brown, E., Brown, M., Bryant, N. P.,
 Buhay, C., Burch, P., Burkett, C., Burrall, K. L., Byrd, N. C.,
 Carron, T. P., Carter, M., Cavazos, S. R., Chacko, J., Chavez, D.,
 Chen, G., Chen, R., Chen, Z., Chowdhury, L., Christopoulos, C.,
 Cleveland, C. D., Cox, C., Coyle, M. D., Dathorne, S. R., David, R.,
 Davila, M. L., Davis, C., Davy-Carroll, L., Dedrich, D. A.,
 Delaney, K. R., Delgado, O., Dem, A. L., Ding, Y., Dinh, H. H.,
 Douthwaite, K. J., Draper, H., Dugan-Rocha, S., Durbin, K. J.,
 Barnhart, C., Edgar, D., Edwards, C. C., Elhaj, C., Escotto, M.,
 Falls, T., Ferraguto, D., Flagg, N., Ford, D., Foster, P., Frantz, P.,
 Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,

Gorrell, J.H., Guevara, M., Gamaral, P., Hale, S., Hamilton, K.,
 Hernandez, C., Harris, K., Hart, M., Havlak, P., Haves, A., Hernandez, J.,
 Hernandez, O., Hodgson, A., Hognes, M., Hollaway, C., Hollins, B.,
 Homai, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.B.,
 Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
 Karlsson, E., Kelly, S., Khan, U., King, L., Korah, J., Kovar, C.,
 Kravovic, J., Kuresh, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
 Li, J., Li, Z., Lichtharge, O., Lien, C., Liu, J., Liu, W., Louised, H.,
 Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
 Masera, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
 Massey, E., Mawney, E., McLeod, P.P., Meador, M., Mei, G., Metzger, M.,
 Mier, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
 Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
 Nguyen, N., Nickerson, E., Nwokenkwo, S., Oguh, M., Okunolu, G.,
 Ogunyemi, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
 Peters, L., Pickens, R., Prims, E., Pu, L.L., Quiles, M., Ren, Y.,
 Rivers, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savary, G.,
 Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sisson, I.,
 Sodergren, E., Sonatke, T., Sparks, A., Stanley, H., Stone, H.,
 Sutton, A., Swatek, A., Taber, P., Tamerisa, A., Tamerisa, K., Yang, H.,
 Tanton, A., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
 Tansley, J., Vasquez, L., Vera, V., Villalón, D., Winsor, R., Wang, Q.,
 Usmani, K., Vasquez, L., Vera, V., Villalón, D., Winsor, R., Wang, Q.,
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Worlington, S.,
 Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K.,
 Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
 Weinstein, G. and Gibbs, R.
 Direct Submission
 2 (bases 1 to 151395)
 Unpublished
 Morley, K.C.
 Direct Submission
 Submitted (24-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 ----- Genome Center -----
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information -----
 Center project name: KAMI
 Center clone name: CH230-43707
 ----- Summary Statistics -----
 Sequencing vector: Plasmid
 Chemistry: Dye-terminator Big Dye 100% of reads
 Assembly program: Phrap, version 0.990329
 Consensus quality: 88196 bases at least Q40
 Consensus quality: 94027 bases at least Q30
 Consensus quality: 97692 bases at least Q20

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 62 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 *
 * 1 1380: contig of 1380 bp in length
 * 1381 1480: gap of unknown length
 * 1481 2902: contig of 1422 bp in length
 * 2903 3002: gap of unknown length
 * 3003 4595: contig of 1593 bp in length
 * 4596 4695: gap of unknown length
 * 4696 6308: contig of 1613 bp in length
 * 6309 6408: gap of unknown length
 * 6409 8011: contig of 1603 bp in length
 * 8012 8111: gap of unknown length
 * 8112 9345: contig of 1234 bp in length
 * 9346 9445: gap of unknown length
 * 9446 10512: contig of 1067 bp in length
 *
 * 10513 10612: gap of unknown length
 * 10613 11823: contig of 1211 bp in length
 * 11824 11923: gap of unknown length
 * 11924 13622: contig of 1699 bp in length
 * 13623 13722: gap of unknown length
 * 13723 14750: contig of 1028 bp in length
 * 14751 14850: gap of unknown length
 * 14851 16300: contig of 1450 bp in length
 * 16301 16400: gap of unknown length
 * 16401 18220: contig of 1820 bp in length
 * 18221 18320: gap of unknown length
 * 18321 20138: contig of 1818 bp in length
 * 20139 20238: gap of unknown length
 * 20239 21480: contig of 1242 bp in length
 * 21481 22983: contig of 1403 bp in length
 * 22984 23083: gap of unknown length
 * 23084 24507: contig of 1424 bp in length
 * 24508 24607: gap of unknown length
 * 24608 26377: contig of 1770 bp in length
 * 26378 26477: gap of unknown length
 * 26478 27917: contig of 1440 bp in length
 * 27918 28017: gap of unknown length
 * 28018 29273: contig of 1256 bp in length
 * 29274 29373: gap of unknown length
 * 29374 31142: contig of 1769 bp in length
 * 31143 31242: gap of unknown length
 * 31243 32285: contig of 1043 bp in length
 * 32286 32385: gap of unknown length
 * 32386 33650: contig of 1265 bp in length
 * 33651 33750: gap of unknown length
 * 33751 36212: contig of 2462 bp in length
 * 36213 36312: gap of unknown length
 * 36313 37975: contig of 1663 bp in length
 * 37976 38075: gap of unknown length
 * 38076 40176: contig of 2101 bp in length
 * 40177 40276: gap of unknown length
 * 40277 42402: contig of 2126 bp in length
 * 42403 42502: gap of unknown length
 * 42503 43849: contig of 1347 bp in length
 * 43850 43949: gap of unknown length
 * 43950 45066: contig of 1137 bp in length
 * 45067 45186: gap of unknown length
 * 45187 46727: contig of 1541 bp in length
 * 46728 46827: gap of unknown length
 * 46828 48234: contig of 1407 bp in length
 * 48235 48334: gap of unknown length
 * 48335 49355: contig of 1621 bp in length
 * 49356 50055: gap of unknown length
 * 50056 52476: contig of 2421 bp in length
 * 52477 52576: gap of unknown length
 * 52577 54907: contig of 2331 bp in length
 * 54908 55007: gap of unknown length
 * 55008 57389: contig of 2382 bp in length
 * 57390 57489: gap of unknown length
 * 57490 60019: contig of 2530 bp in length
 * 60020 60119: gap of unknown length
 * 60120 62219: contig of 2100 bp in length
 * 62220 62319: gap of unknown length
 * 62320 64250: contig of 1931 bp in length
 * 64251 64350: gap of unknown length
 * 64351 65989: contig of 1639 bp in length
 * 65990 66089: gap of unknown length
 * 66090 68073: contig of 1984 bp in length
 * 68074 68173: gap of unknown length
 * 68174 70180: contig of 2007 bp in length
 * 70181 70280: gap of unknown length
 * 70281 72186: contig of 1906 bp in length
 * 72187 72286: gap of unknown length
 * 72287 74046: contig of 1760 bp in length
 * 74047 74146: gap of unknown length
 * 74147 76610: contig of 2464 bp in length
 * 76611 76710: gap of unknown length

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

76711 79026: contig of 2316 bp in length
 79027 79126: gap of unknown length
 79127 81402: contig of 2276 bp in length
 81403 81403: gap of unknown length
 85002 85002: contig of 3507 bp in length
 85010 85109: gap of unknown length
 85110 87545: contig of 2436 bp in length
 87546 87546: gap of unknown length
 87646 89759: contig of 2114 bp in length
 89760 89859: gap of unknown length
 92889 92888: contig of 3029 bp in length
 92889 92988: gap of unknown length
 92989 96176: contig of 3188 bp in length
 96177 96277: gap of unknown length
 96277 99868: contig of 3592 bp in length
 99869 99969: gap of unknown length
 104701 104800: gap of unknown length
 108393 108392: contig of 3592 bp in length
 108393 108492: gap of unknown length
 108493 111786: contig of 3294 bp in length
 111787 111886: gap of unknown length
 111887 116187: contig of 4301 bp in length
 116188 116287: gap of unknown length
 116288 119502: contig of 3215 bp in length
 119503 123632: contig of 4030 bp in length
 123633 123732: gap of unknown length

Query Match 91.6%; Score 17.4; DB 2; Length 151395;
 Best Local Similarity 94.7%; Pred. No. 1.5e+02;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CGCTGATTCGCTGCTG 19
 73184 CTCTGATTCGCTGCTG 73202

RESULT 11
 AC119463/c
 LOCUS
 DEFINITION Rattus norvegicus clone CH230-22502, *** SEQUENCING IN PROGRESS
 AC119463
 AC119463.3 GI:21746638
 HG: HMG PHASE1
 VERSION
 KEYWORDS
 SOURCE Norway rat
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 1 (bases 1 to 168369)
 Muszy,D.M., Adams,C., Adio-Oduola,B., Ali-oshan,F.R., Allen,C., Alsbrooks,S.L., Amaratunga,H.C., Are,J.R., Ayale,M., Banks,T., Barbata,U., Benton,J., Bimaga,K., Blankenburg,K., Bonin,D., Bouck,J., Bowle,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,W.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Demu,A.B., Ding,Y., Dinch,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.T., Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hognes,M., Hollaway,C., Hollins,B., Homsi,F., Howard,S., Huber,J., Hulik,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudan,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,

TITLE JOURNAL REFERENCE AUTHORS JOURNAL

REFERENCE AUTHORS TITLE JOURNAL COMMENT

Li,J., Li,Z., Lichtarge,O., Lien,C., Liu,J., Liu,W., Loulseged,H., Lozano,R.J., Lu,X., Lucier,A., Luckler,R., Luna,R., Ma,J., Maheshwari,M., Mapa,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawliny,E., McLeod,M.P., Meador,M., Mei,G., Metzger,M., Miner,G., Miner,Z., Mitchell,T., Mohabhat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenwo,S., Oguth,M., Okunodu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rivers,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoshitari,N., Sisson,I., Sodergren,E., Sonlike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Syarik,A., Taboz,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usami,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q., Wang,S., Ward-More,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczek,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstein,G. and Gibbs,R.
 Direct Submission
 Unpublished
 2 (bases 1 to 168369)
 Morley,K.C.
 Direct Submission
 Submitted (27-APR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 168369)
 Morley,K.C.
 Direct Submission
 Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 On Jul 14, 2002 this sequence version replaced gi:20387276.
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GUTZ
 Center clone name: CH230-22502
 ----- Summary Statistics
 Sequencing vector: Plasmid
 Chemistry: Dye-terminator Big Dye 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 104715 bases at least Q40
 Consensus quality: 111142 bases at least Q30
 Consensus quality: 116085 bases at least Q20

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a "working draft" sequence. It currently
 * consists of 74 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 1 1052: contig of 1052 bp in length
 * 1053 1152: gap of unknown length
 * 1153 2488: contig of 1336 bp in length
 * 2489 2589: gap of unknown length
 * 2589 4101: contig of 1513 bp in length
 * 4102 4201: gap of unknown length
 * 4202 5283: contig of 1082 bp in length
 * 5284 5383: gap of unknown length
 * 5384 6393: contig of 1010 bp in length
 * 6394 6493: gap of unknown length
 * 6494 8259: contig of 1766 bp in length
 * 8260 8359: gap of unknown length
 * 8360 9702: contig of 1343 bp in length

```

* 9703 9802: gap of unknown length
* 9803 1137: contig of 1335 bp in length
* 1138 1137: gap of unknown length
* 1138 12257: contig of 1020 bp in length
* 12358 12357: gap of unknown length
* 12358 13955: contig of 1598 bp in length
* 13956 14055: gap of unknown length
* 14056 15937: contig of 1882 bp in length
* 15938 16037: gap of unknown length
* 16038 17706: contig of 1669 bp in length
* 17707 17806: gap of unknown length
* 17807 19352: contig of 1546 bp in length
* 19353 19452: gap of unknown length
* 19453 20883: contig of 1431 bp in length
* 20884 20983: gap of unknown length
* 20984 22048: contig of 1065 bp in length
* 22049-22049: gap of unknown length
* 22149 23896: contig of 1748 bp in length
* 23897 23996: gap of unknown length
* 23997 25356: contig of 1360 bp in length
* 25357 25456: gap of unknown length
* 25457 26764: contig of 1308 bp in length
* 26765 26864: gap of unknown length
* 26865 27982: contig of 1118 bp in length
* 27983 28082: gap of unknown length
* 28083 29140: contig of 1058 bp in length
* 29141 30671: contig of 1431 bp in length
* 30672 30772: gap of unknown length
* 30772 32820: contig of 2049 bp in length
* 32821 32920: gap of unknown length
* 32921 34405: contig of 1485 bp in length
* 34406 34505: gap of unknown length
* 34506 36497: contig of 1992 bp in length
* 36498 36597: gap of unknown length
* 36598 37336: contig of 1033 bp in length
* 37337 37736: gap of unknown length
* 37737 39316: contig of 1580 bp in length
* 39317 39416: gap of unknown length
* 39417 40439: contig of 1023 bp in length
* 40440 40539: gap of unknown length
* 40540 41568: contig of 1023 bp in length
* 41569 41668: gap of unknown length
* 41669 42719: contig of 1051 bp in length
* 42720 42819: gap of unknown length
* 42820 44594: contig of 1775 bp in length
* 44595 44694: gap of unknown length
* 44695 45911: contig of 1217 bp in length
* 45912 46011: gap of unknown length
* 46012 47510: contig of 1499 bp in length
* 47511 47610: gap of unknown length
* 47611 48721: contig of 1111 bp in length
* 48722 48821: gap of unknown length
* 48822 50922: contig of 2101 bp in length
* 50923 51022: gap of unknown length
* 51023 52233: contig of 1211 bp in length
* 52234 52333: gap of unknown length
* 52334 54084: contig of 1751 bp in length
* 54085 54184: gap of unknown length
* 54185 55537: contig of 1353 bp in length
* 55538 55637: gap of unknown length
* 55638 57248: contig of 1611 bp in length
* 57249 57348: gap of unknown length
* 57349 59510: contig of 2162 bp in length
* 59511 59610: gap of unknown length
* 59611 60612: contig of 1002 bp in length
* 60613 60712: gap of unknown length
* 60713 62445: contig of 1733 bp in length
* 62446 62545: gap of unknown length
* 62546 64053: contig of 1508 bp in length
* 64054 64153: gap of unknown length
* 64154 66613: contig of 2460 bp in length
* 66614 66713: gap of unknown length

```

```

* 66714 66864: contig of 1951 bp in length
* 66865 68764: gap of unknown length
* 68765 70129: contig of 1365 bp in length
* 70130 70229: gap of unknown length
* 70230 72519: contig of 2290 bp in length
* 72520 72619: gap of unknown length
* 72620 74651: contig of 2032 bp in length
* 74652 74751: gap of unknown length
* 74752 76057: contig of 1306 bp in length
* 76058 76157: gap of unknown length
* 76158 78182: contig of 2025 bp in length
* 78183 78282: gap of unknown length
* 78283 79881: contig of 1593 bp in length
* 79882 82071: contig of 2090 bp in length
* 82072 82171: gap of unknown length
* 82172 83559: contig of 1388 bp in length
* 83560 83659: gap of unknown length
* 83660 85035: contig of 1376 bp in length
* 85036 85135: gap of unknown length
* 85136 87276: contig of 2141 bp in length

Query Match 91.6%; Score 17.4; DB 2; length 168369;
Best Local Similarity 94.7%; Pred. No. 1.5e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGCTCGATTCGCTCTG 19
Db 4044 CGCTCGATTCGCTCTG 4026

RESULT 12
AC022377
LOCUS
DEFINITION Homo sapiens chromosome 3, WORKING DRAFT SEQUENCE, 6 unordered
pieces.
ACCESSION AC022377
VERSION AC022377.2 GI:10305054
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 176054)
Lin.W., Zhang,C., Dong,H., Wan,M., Xu,S., Gu,W., Tu,Y., Jia,J.,
Wu,C., Lu,G., Zhong,M., Zhou,Y., Ren,S., Fu,G., Chen,Z. and
Huang,M.
Direct Submission
Submitted (03-FEB-2000) Genomic Dept., Chinese National Human
Genome Center at Shanghai, 315 Guo Shou Jing Road, Shanghai,
Shanghai 201203, P. R. China
On Sep 26, 2000 this sequence version replaced gi:1682623.
-----Genome Center-----
Center:Beijing Center
Center code:Beijing
Website:http://hgsc.igtp.ac.cn
http://www.genomics.org.cn
Contact:hgsc@igtp.ac.cn.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 8523: contig of 8523 bp in length
* 8524 21709: contig of 13186 bp in length
* 21710 39761: contig of 18052 bp in length
* 39762 42545: contig of 2784 bp in length

```


/note="assembly_fragment:01782
fragment_chain:1
clone_end:17
vector_side:right"
BASE COUNT 46064 a 48225 c 48472 g 48190 t 1225 others
ORIGIN

Query Match 91.6%; Score 17.4; DB 2; Length 192180;
Best Local Similarity 94.7%; Pred. No. 1.5e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGCTCGATTCTCTGCTG 19
|||
Db 27976 CGCTCGATTCTCTGCGCTG 27994

Search completed: June 21, 2003, 19:56:22
Job time : 997 secs